

QC

1 =

A85

V. 11

1958

SCIENCE

PEA

Digitized by the Internet Archive
in 2024

AUSTRALIAN JOURNAL OF PHYSICS

VOLUME II

MELBOURNE
1958

AUSTRALIAN JOURNAL OF PHYSICS

Published by the Commonwealth Scientific and Industrial Research Organization. Volumes 1 to 5 of the Australian Journal of Physics and the Australian Journal of Chemistry issued as the Australian Journal of Scientific Research, Series A : Physical Sciences. Issued quarterly, £2 per annum.

BOARD OF STANDARDS

The Board of Standards for this Journal is appointed by the Commonwealth Scientific and Industrial Research Organization and the Australian Academy of Science and consists of Professor J. G. Wood (Chairman), Dr. N. S. Noble (Editor), Professor J. S. Anderson, Professor Sir Macfarlane Burnet, Professor Sir Leslie Martin, and Professor W. P. Rogers.

ADVISORY COMMITTEE

Acceptance of papers for this Journal is in the hands of an Advisory Committee appointed by the Board of Standards in consultation with the Institute of Physics (Australian Branch) and consisting of Dr. N. S. Noble (Chairman and Editor), Dr. G. H. Briggs, Professor L. G. H. Huxley, Professor Sir Leslie Martin, and Professor H. C. Webster.

All enquiries and manuscripts should be forwarded to :

The Editor,
Australian Journal of Physics,
Commonwealth Scientific and Industrial Research Organization,
314 Albert Street, East Melbourne, C.2, Victoria

CONTENTS

NUMBER 1, MARCH 1958

	PAGE
Coupling of Nuclear Spins in Molecules. By D. W. Posener	1
The Townsend Ionization Coefficients in Crossed Electric and Magnetic Fields. By H. A. Blevin and S. C. Haydon	18
High Resolution Cinematography of the Solar Photosphere. By R. E. Loughhead and V. R. Burgess	35
The Design of Photographic Objectives of the Triplet Family. I. The Design of the Triplet Type III Objective. By F. D. Cruickshank ..	41
Further Observations of Radio Emission from the Planet Jupiter. By F. F. Gardner and C. A. Shain	55
Gain Measurements of Large Aerials used in Interferometer and Cross-type Radio Telescopes. By A. G. Little	70
Anomalies in Ionosonde Records due to Travelling Ionospheric Disturbances. By L. H. Heisler	79
Travelling Ionospheric Disturbances in the <i>F</i> Region. By G. H. Munro ..	91
The 1956 Phoenicid Meteor Shower. By A. A. Weiss	113
Electromagnetic Propagation in an Almost Homogeneous Medium. By V. W. Bolie	118

Short Communications

Low Latitude Reflections from the Aurora Australis. By T. J. Seed and C. D. Ellyett	126
The Distribution of Flare Heights as Derived from Limb Flares. By R. G. Giovanelli and Marie K. McCabe	130
Decay Time of the Luminescence of a Zinc Sulphide Neutron Detector for Neutron and γ -Ray Excitation. By G. M. Bailey and J. R. Prescott	135
Coulomb Wave Functions. By A. Learner and B. A. Robson	138

NUMBER 2, JUNE 1958

PAGE

Methods for Numerical Calculations with the Type I Counter. By P. G. Guest	143
The General Relationships between the Elastic Constants of Isotropic Materials in n Dimensions. By N. W. Tschoegl	154
The Ionization of Liquid Argon by α -Particles. By F. D. Stacey ..	158
Thermal Structures in the Lowest Layers of the Atmosphere. By R. J. Taylor	168
The Wilson Effect in Sunspots. By R. E. Loughhead and R. J. Bray	177
The Lifetime of Sunspot Penumbra Filaments. By R. J. Bray and R. E. Loughhead	185
The Flare-surge Event. By R. G. Giovanelli and Marie K. McCabe ..	191
Polarization Measurements of the Three Spectral Types of Solar Radio Burst. By M. Komesaroff	201
Evidence of Echoes in the Solar Corona from a New Type of Burst. By J. A. Roberts	215
On the Galvanomagnetic, Thermomagnetic, and Thermoelectric Effects in Isotropic Metals and Semiconductors. By E. J. Moore	235
On the Cylindrical Probe Method of Measuring Thermal Conductivity with Special Reference to Soils. I. Extension of Theory and Discussion of Probe Characteristics. By D. A. de Vries and A. J. Peck	255

Short Communications

Lunar Tides in E_{2s} at Brisbane. By A. D. Gazzard	272
Photoneutrons from Natural Magnesium. By B. M. Spicer, F. R. Allum, J. E. E. Baglin, and H. H. Thies	273
A Suggested Improvement to the C.W. Technique for Measurement of Meteor Velocities. By J. S. Mainstone, W. G. Elford, and A. A. Weiss	277

NUMBER 3, SEPTEMBER 1958

	PAGE
Ether and Relativity. By G. Builder	279
The Giant Resonance of Photodisintegration of Tantalum. By B. M. Spicer, H. H. Thies, J. E. Baglin, and F. R. Allum	298
The Fluctuating Field Ferromagnet at Low Temperatures. By F. D. Stacey	310
On the Seismological Aspects of the 1954 Hydrogen Bomb Explosions. By T. N. Burke-Gaffney and K. E. Bullen	318
A Study of "Spread- <i>F</i> " Ionospheric Echoes at Night at Brisbane. IV. Range Spreading. By H. C. Webster	322
Solar Brightness Distribution at a Wavelength of 60 Centimetres. II. Localized Radio Bright Regions. By G. Swarup and R. Parthasarathy	338
Flare-Puffs as a Cause of Type III Radio Bursts. By R. G. Giovanelli ..	350
Optical Observations of the Solar Disturbances causing Type II Radio Bursts. By R. G. Giovanelli and J. A. Roberts	353
A Catalogue of Radio Sources between Declinations $+10^\circ$ and -20° . By B. Y. Mills, O. B. Slee, and E. R. Hill	360
On the Radio Emission of Hydrogen Nebulae. By C. M. Wade ..	388
An Investigation of the Strong Radio Sources in Centaurus, Fornax, and Puppis. By K. V. Sheridan	400
On the Cylindrical Probe Method of Measuring Thermal Conductivity with Special Reference to Soils. II. Analysis of Moisture Effects. By D. A. de Vries and A. J. Peck	409
Electromagnetic Radiation from Electrons Rotating in an Ionized Medium under the Action of a Uniform Magnetic Field. By R. Q. Twiss and J. A. Roberts	424

Short Communications

Ferromagnetic Exchange between Coupled Pairs of Electrons. By F. D. Stacey	447
The Soft X-ray L_{23} Emission Spectrum of Magnesium from Solid and Evaporated Targets. By R. S. Crisp	449
Freezing Nucleus Measurements in January 1957. By E. G. Bowen ..	452

NUMBER 4, DECEMBER 1958

	PAGE
The Constancy of the Velocity of Light. By G. Builder	457
Electron Excitation of Collective Nuclear Transitions. By L. J. Tassie ..	481
The Photodisintegration of Nuclei with Z between 9 and 30. By B. M. Spicer	490
Excited States of ${}^8\text{Be}$ from the ${}^7\text{Li}(d,n){}^8\text{Be}$ Reaction. By R. H. Spear	502
Observations of Changes in the Photospheric Granules. By R. J. Bray and R. E. Loughhead	507
The Radio Emission from Centaurus-A and Fornax-A. By C. A. Shain ..	517
A Pencil-beam Survey of the Galactic Plane at 3.5 m. By E. R. Hill, O. B. Slee, and B. Y. Mills	530
Radio Emission from the Vela-Puppis Region. By H. Rishbeth ..	550
Radiation Transfer and the Possibility of Negative Absorption in Radio Astronomy. By R. Q. Twiss	564

Short Communications

A Search for Radio Emission at 3.5 m from the Local Supergalaxy. By E. R. Hill	580
The Clock Paradox in Special Relativity. By H. Jeffreys	583
The Clock Paradox in Relativity. By E. F. Fahy	586
Variations in Ionospheric F -region Characteristics. By N. M. Brice ..	587
Approximations for the Electron Density in Meteor Trails. By A. A. Weiss	591
Index to Volume 11	595

COUPLING OF NUCLEAR SPINS IN MOLECULES

By D. W. POSENER*

[Manuscript received June 19, 1957]

Summary

The theory of quadrupole, magnetic dipole, and dipole-dipole interactions of nuclear spins with molecular rotation is generalized for any number of nuclear spins in any free molecule which has no resultant electronic angular momentum.

Calculation of the matrix elements of the Hamiltonian is discussed in detail, and the cases of one and two nuclear spins are dealt with explicitly.

I. INTRODUCTION

First-order matrix elements for the coupling of two similar quadrupolar nuclei in a molecule have been derived by Foley (1947) for the diatomic case, and subsequently by Myers and Gwinn (1952) and Robinson and Cornwell (1953) for more general molecules. Cases of grossly unequal coupling have been considered by Bardeen and Townes (1948*a*, 1948*b*) and by Townes and Schawlow (1955, Section 6-6), although the appropriate matrix elements have not been given explicitly.

Bersohn (1950) has given the matrix elements for the quadrupolar coupling of three nuclei, and his methods, which are applicable to any number of nuclei, are described in his thesis (Bersohn 1949).

Recent advances in high-resolution microwave spectroscopy have drawn more attention to the small magnetic interactions of nuclei (White 1955). Gunther-Mohr, Townes, and Van Vleck (1954) and Gordon (1955) have described the coupling of three hydrogen spins and a quadrupolar nucleus (nitrogen) in ammonia; Okaya (1956) has considered C_{2v} molecules with two $\frac{1}{2}\hbar$ off-axis spins and one axial quadrupolar nucleus; and Herrmann (1956) has treated the experimental problems of ND_3 using some theoretical results derived by Hadley (1955), whose work has not been accessible to the present writer. Although these cases are of increasing complexity, symmetry properties of the particular molecular types so far considered help to provide some simplification of the problems, and the works quoted cannot be easily generalized.

In the following account we will deal with an arbitrary free molecule with no resultant electronic angular momentum, but containing n nuclei each possessing spin, and we will give the complete matrix elements of the major interactions. The results are derived for quadrupolar spins, but can be easily specialized to cases where some of the spins are $\frac{1}{2}\hbar$ or zero.

In most problems of practical interest only a very few of the terms given need to be considered because some terms contribute to the energy only in

* Division of Electrotechnology, C.S.I.R.O., University Grounds, Chippendale, N.S.W.

higher order of approximation, while in many cases some of the matrix elements vanish identically because of symmetry properties of the molecule. Because of their wide variety symmetry considerations peculiar to particular types of molecules will not be specially discussed here, although they will usually be of great importance in further simplifying the calculations in a given problem. We refer to some of the papers already quoted for discussions of this kind.

The results for one and for two nuclei will be given explicitly as simple applications of the general theory.

II. THE HAMILTONIAN

We assume the molecule to be subject to no external fields and to have no resultant electronic orbital or spin angular momentum, so that electron spins are paired and their effects may usually be neglected (as demonstrated, for instance, by Gunther-Mohr, Townes, and Van Vleck (1954)). In the rigid rotor approximation the Hamiltonian for rotation and for the major spin hyperfine interactions is then (Van Vleck 1951; Gunther-Mohr, Townes, and Van Vleck 1954):

$$\begin{aligned}
 H = & \sum_g \bar{G}_g (J_g - L_g)^2 + \frac{e\mu_N}{c} \sum_k \sum_i r_{ik}^{-3} (\mathbf{r}_i - \mathbf{r}_k) \times \left[\mathbf{v}_i - \left(1 + \frac{Z_k M_P}{g_k M_k} \right) \mathbf{v}_k \right] \cdot g_k \mathbf{I}_k \\
 & - \frac{e\mu_N}{c} \sum_{l \neq k} \sum_i Z_i r_{kl}^{-3} (\mathbf{r}_k - \mathbf{r}_l) \times \left[\left(1 + \frac{Z_k M_P}{g_k M_k} \right) \mathbf{v}_k - \mathbf{v}_l \right] \cdot g_k \mathbf{I}_k \\
 & + \mu_N^2 \sum_k \sum_{l > k} r_{kl}^{-3} \{ g_k \mathbf{I}_k \cdot g_l \mathbf{I}_l - 3 r_{kl}^{-2} [g_k \mathbf{I}_k \cdot (\mathbf{r}_k - \mathbf{r}_l)] [g_l \mathbf{I}_l \cdot (\mathbf{r}_k - \mathbf{r}_l)] \} \\
 & + \sum_k \frac{1}{12} \left(2 \frac{\partial^2 V_k}{\partial z_k^2} - \frac{\partial^2 V_k}{\partial x_k^2} - \frac{\partial^2 V_k}{\partial y_k^2} \right) \int \rho_k (2z_k'^2 - x_k'^2 - y_k'^2) dv_k. \quad \dots \quad (1)
 \end{aligned}$$

where \mathbf{J} is the total angular momentum exclusive of nuclear spin, \mathbf{L} is the electronic orbital angular momentum, and we measure angular momentum in units of \hbar . The \bar{G}_g are rotation constants, with g referring to principal inertial axes x , y , and z fixed in the molecule. The sum over i is over the electrons, whose charges are $-e$ and whose positions and velocities are given by \mathbf{r}_i and \mathbf{v}_i respectively, referred to the molecular centre of mass. Indices k and l are used similarly for the n nuclei, which have magnetic moments $g_k \mu_N$, masses M_k , charges Z_k , spins \mathbf{I}_k , and nuclear charge densities ρ_k . The proton mass is M_P , μ_N is the nuclear magneton, c is the velocity of light, and r_{ik} is an abbreviation for $|\mathbf{r}_i - \mathbf{r}_k|$. The electrostatic potential at the k th nucleus due to all molecular charges outside the nuclear region is V_k , and x'_k , y'_k , and z'_k are coordinates fixed in the k th nucleus with the z'_k direction along the nuclear axis of symmetry.

The first term in the Hamiltonian is the energy of rigid rotation, the second term represents the energy of interaction of nuclear magnetic dipoles with currents due to electron motion, the third term is the energy of interaction of nuclear dipoles with currents due to nuclear motion (molecular rotation), the fourth term is the magnetic dipole-dipole interaction of the nuclei with one another, and the last term is the energy of nuclear quadrupolar interaction with the molecular electrostatic fields.

Although the non-rotating molecule possesses zero total electron orbital angular momentum, rotational interactions excite higher electronic states. When perturbation theory is used to take the more important interactions of these states into account (see, for example, Gunther-Mohr, Townes, and Van Vleck 1954), the effective Hamiltonian for the nuclear spin interactions can be written in the form:

$$\begin{aligned}
 H = & \sum_k [a'_k(\bar{I}_{kx}J_x + \bar{I}_{ky}J_y + \bar{I}_{kz}J_z) + a_k(2\bar{I}_{kz}J_z - \bar{I}_{kx}J_x - \bar{I}_{ky}J_y) \\
 & + b_k(\bar{I}_{kx}J_x - \bar{I}_{ky}J_y) + c_k(\bar{I}_{kx}J_y + \bar{I}_{ky}J_x) + d_k(\bar{I}_{kx}J_z + \bar{I}_{kz}J_x) \\
 & + e_k(\bar{I}_{ky}J_z + \bar{I}_{kz}J_y)] + \sum_k \sum_{l>k} [\alpha_S^{(kl)}(2\bar{I}_{kz}\bar{I}_{lz} - \bar{I}_{kx}\bar{I}_{lx} - \bar{I}_{ky}\bar{I}_{ly}) \\
 & + \beta_S^{(kl)}(\bar{I}_{kx}\bar{I}_{lx} - \bar{I}_{ky}\bar{I}_{ly}) + \gamma_S^{(kl)}(\bar{I}_{kx}\bar{I}_{ly} + \bar{I}_{ky}\bar{I}_{lx}) \\
 & + \delta_S^{(kl)}(\bar{I}_{kx}\bar{I}_{lz} + \bar{I}_{kz}\bar{I}_{lx}) + \varepsilon_S^{(kl)}(\bar{I}_{ky}\bar{I}_{lz} + \bar{I}_{kz}\bar{I}_{ly})] \\
 & + \sum_k [\alpha_k(2\bar{I}_{kz}^2 - \bar{I}_{kx}^2 - \bar{I}_{ky}^2) + \beta_k(\bar{I}_{kx}^2 - \bar{I}_{ky}^2) + \gamma_k(\bar{I}_{kx}\bar{I}_{ky} + \bar{I}_{ky}\bar{I}_{kx}) \\
 & + \delta_k(\bar{I}_{kx}\bar{I}_{kz} + \bar{I}_{kz}\bar{I}_{kx}) + \varepsilon_k(\bar{I}_{ky}\bar{I}_{kz} + \bar{I}_{kz}\bar{I}_{ky})]. \quad \dots\dots\dots (2)
 \end{aligned}$$

For convenience in calculation everything in equation (2) is referred to the molecule-fixed system of axes, and for consistent commutation relations the spin angular momenta \mathbf{I}_k have been replaced by their reverses $\bar{\mathbf{I}}_k = -\mathbf{I}_k$ (Van Vleck 1951). The coefficients in equation (2) are:

$$\begin{aligned}
 a'_k &= \frac{1}{3}(M_{xx}^{(k)} + M_{yy}^{(k)} + M_{zz}^{(k)}), \\
 a_k &= \frac{1}{6}(2M_{zz}^{(k)} - M_{xx}^{(k)} - M_{yy}^{(k)}), \\
 b_k &= \frac{1}{2}(M_{xx}^{(k)} - M_{yy}^{(k)}), \\
 c_k &= \frac{1}{2}(M_{xy}^{(k)} + M_{yx}^{(k)}), \\
 d_k &= \frac{1}{2}(M_{xz}^{(k)} + M_{zx}^{(k)}), \\
 e_k &= \frac{1}{2}(M_{yz}^{(k)} + M_{zy}^{(k)}), \\
 \alpha_S^{(kl)} &= -\frac{1}{2}\mu_N^2 g_k g_l [2(r_{kl})_z^2 - (r_{kl})_x^2 - (r_{kl})_y^2] / r_{kl}^5, \\
 \beta_S^{(kl)} &= -\frac{3}{2}\mu_N^2 g_k g_l [(r_{kl})_x^2 - (r_{kl})_y^2] / r_{kl}^5, \\
 \gamma_S^{(kl)} &= -3\mu_N^2 g_k g_l (r_{kl})_x (r_{kl})_y / r_{kl}^5, \\
 \delta_S^{(kl)} &= -3\mu_N^2 g_k g_l (r_{kl})_x (r_{kl})_z / r_{kl}^5, \\
 \varepsilon_S^{(kl)} &= -3\mu_N^2 g_k g_l (r_{kl})_y (r_{kl})_z / r_{kl}^5, \\
 \alpha_k &= \frac{eQ_k}{12I_k(2I_k-1)}(2V_{zz}^{(k)} - V_{xx}^{(k)} - V_{yy}^{(k)}), \\
 \beta_k &= \frac{eQ_k}{4I_k(2I_k-1)}(V_{xx}^{(k)} - V_{yy}^{(k)}), \\
 \gamma_k &= \frac{eQ_k}{2I_k(2I_k-1)}V_{xy}^{(k)}, \\
 \delta_k &= \frac{eQ_k}{2I_k(2I_k-1)}V_{xz}^{(k)}, \\
 \varepsilon_k &= \frac{eQ_k}{2I_k(2I_k-1)}V_{yz}^{(k)}, \\
 eQ_k &= \int \rho_k(2z^2 - x^2 - y^2)dv_k, \\
 V_{gg'}^{(k)} &= \frac{\partial^2 V_k}{\partial g \partial g'},
 \end{aligned}$$

$$\begin{aligned}
M_{gg'}^{(k)} = & \frac{2e\mu_N g_k \bar{G}_g}{\hbar c} \left\langle \left(1 + \frac{Z_k M_P}{g_k M_k} \right) \sum_i r_{ik}^{-3} [\delta_{gg'} (\mathbf{r}_i - \mathbf{r}_k) \cdot \mathbf{r}_k - (\mathbf{r}_i - \mathbf{r}_k)_g (\mathbf{r}_k)_g] \right. \\
& + \sum_{l \neq k} Z_l r_{kl}^{-3} \left\{ \delta_{gg'} (\mathbf{r}_k - \mathbf{r}_l) \cdot \left[\left(1 + \frac{Z_k M_P}{g_k M_k} \right) \mathbf{r}_k - \mathbf{r}_l \right] \right. \\
& \quad \left. \left. - (\mathbf{r}_k - \mathbf{r}_l)_g \left[\left(1 + \frac{Z_k M_P}{g_k M_k} \right) \mathbf{r}_k - \mathbf{r}_l \right]_g \right\} \right\rangle \\
& + \frac{2e\mu_N g_k \bar{G}_g}{c} \sum_p \frac{(0 | L_g | p)(p | \Pi_{g'}^{(k)} | 0) + (0 | \Pi_{g'}^{(k)} | p)(p | L_g | 0)}{E_0 - E_p}, \\
\Pi_g^{(k)} = & \sum_i r_{ik}^{-3} [(\mathbf{r}_i - \mathbf{r}_k) \times \mathbf{v}_i]_g, \quad \dots \dots \dots (3)
\end{aligned}$$

in which $\delta_{gg'}$ is the Kronecker delta, p specifies an excited electronic state of energy E_p , and the last sum indicated in $M_{gg'}^{(k)}$ is to be taken over all values of p except $p=0$. $(r_{kl})_g$ is short for $(\mathbf{r}_k - \mathbf{r}_l)_g$. The functions of $(r_{kl})_g$ and of $V_{gg'}^{(k)}$, together with the other averages shown explicitly here and later, are to be averaged over the ground electronic and vibrational state of the molecule.

In most practical cases a large number of these coefficients vanish because of molecular symmetry, or can be neglected since they contribute to the energy only in higher order.

III. ADDITION OF ANGULAR MOMENTA

Bersohn (1949), using the tensor procedure of Racah (1942), has discussed a general method for calculating the matrix elements for the quadrupole coupling of a number of nuclei in a molecule. Condon and Shortley (1953) (hereafter referred to as TAS), following Guttinger and Pauli (1931), have derived the matrix elements of two commuting angular momenta. This section will show that the extension of the methods of TAS to any number of commuting angular momenta is not as formidable as it might first appear, and the results are in some cases rather simpler to use than those of Bersohn.

Since the problem of addition of a number of commuting angular momenta is rather general the notation of TAS will be followed as far as possible, with suitable generalizations and with some of the functions introduced by Van Vleck (1951), and as a basis for our subsequent discussion we shall start by reviewing the appropriate results given in TAS.

(a) Review of TAS Results

In a representation which diagonalizes the square of the total angular momentum \mathbf{J} and its component J_z along some space-fixed Z -axis, the matrix elements of the components of \mathbf{J} are off-diagonal in the quantum number m only, and are given by TAS, Section 2³:

$$\left. \begin{aligned}
(\alpha j m | J_x | \alpha j m \pm 1) &= \frac{1}{2} f(j, \pm m), \\
(\alpha j m | J_y | \alpha j m \pm 1) &= \pm \frac{1}{2} i f(j, \pm m), \\
(\alpha j m | J_z | \alpha j m) &= m,
\end{aligned} \right\} \dots \dots \dots (4)$$

where

$$f(j, m) = \sqrt{\{j(j-m)(j+m+1)\}}, \quad \dots \dots \dots (5)$$

$i = \sqrt{-1}$, and we have used α to represent the totality of unspecified quantum numbers.

If \mathbf{T} is a vector operator which obeys the commutation rule TAS 8³¹, the dependence on m of the components of \mathbf{T} is given by TAS 9³¹¹:

$$\left. \begin{aligned} (\alpha jm | T_x | \alpha' j+1 m \pm 1) &= \mp \frac{1}{2} (\alpha j; T; \alpha' j+1) g(j, \mp m-2), \\ (\alpha jm | T_x | \alpha' j m \pm 1) &= \frac{1}{2} (\alpha j; T; \alpha' j) f(j, \pm m), \\ (\alpha jm | T_x | \alpha' j-1 m \pm 1) &= \pm \frac{1}{2} (\alpha j; T; \alpha' j-1) g(j, \pm m), \\ (\alpha jm | T_y | \alpha' j' m') &= \pm i (\alpha jm | T_x | \alpha' j' m'), \\ (\alpha jm | T_z | \alpha' j+1 m) &= (\alpha j; T; \alpha' j+1) \sqrt{\{(j+1)^2 - m^2\}}, \\ (\alpha jm | T_z | \alpha' j m) &= (\alpha j; T; \alpha' j) m, \\ (\alpha jm | T_z | \alpha' j-1 m) &= (\alpha j; T; \alpha' j-1) \sqrt{\{j^2 - m^2\}}, \end{aligned} \right\} \dots (6)$$

where

$$g(j, m) = \sqrt{\{(j-m)(j-m-1)\}}, \dots (7)$$

and $(\alpha j; T; \alpha' j')$ is independent of m . The equations (6) show that the matrix of \mathbf{T} can be factored into submatrices, one of which contains the whole dependence on the quantum number m .

If

$$\mathbf{J}_1 + \mathbf{J}_2 = \mathbf{J}, \dots (8)$$

and \mathbf{J}_1 and \mathbf{J}_2 commute, TAS 10³ shows that the dependence on m of \mathbf{J}_1 and \mathbf{J}_2 is given by the equations (6), and the submatrices $(j; \mathbf{J}_1; j')$ and $(j; \mathbf{J}_2; j')$ are as shown in Table 1, in which

$$\left. \begin{aligned} \varphi(j_1, j) &= - \frac{\sqrt{\{P(j_1, j)Q(j_1, j-1)\}}}{j\sqrt{\{(2j-1)(2j+1)\}}}, \\ \theta(j_1, j) &= \frac{R(j_1, j)}{2j(j+1)}, \end{aligned} \right\} \dots (9)$$

and

$$\left. \begin{aligned} P(j_1, j) &= (j-j_2+j_1)(j+j_2+j_1+1), \\ Q(j_1, j) &= (j_2+j_1-j)(j+j_2-j_1+1), \\ R(j_1, j) &= j(j+1)-j_2(j_2+1)+j_1(j_1+1). \end{aligned} \right\} \dots (10)$$

Functions of (j_2, j) are obtained by interchanging j_1 and j_2 throughout. Also,

$$\varphi(j_2, j) = \varphi(j_1, j). \dots (11)$$

Here and subsequently we show only two variables (quantum numbers) explicitly since the third is always defined by the coupling scheme (such as equations (8) and (15)); we will use s and t to denote a general pair of such quantum numbers.

For brevity we frequently omit diagonal quantum numbers from the matrix elements.

If \mathbf{P} is any vector operator which commutes with \mathbf{J}_1 , but obeys TAS 8³¹ with respect to \mathbf{J} , then by TAS 11³ the dependence of the components of \mathbf{P} on m is given by the equations (6), and the dependence of $(\alpha j; \mathbf{P}; \alpha' j')$ on j may also be factored out:

$$(\alpha j_2 j; \mathbf{P}; \alpha' j_2 j') = (\alpha j_2; \mathbf{P}; \alpha' j_2) (j_2 j | j_2 j'), \dots (12)$$

where the submatrix $(j_2 j | j'_2 j')$ is given by Table 2 with s and t replaced by j_2 and j . It will be noticed that $(st | s' t')$ is a Hermitian matrix, and this property of the matrices will be used later in other results when for conciseness not all

TABLE 1
THE MATRICES $(j; J_1; j')$ AND $(j; J_2; j')$

j'	$j+1$	j	$j-1$
$(j; J_1; j')$	$-\frac{1}{2}\varphi(j_1, j+1)$	$\theta(j_1, j)$	$-\frac{1}{2}\varphi(j_1, j)$
$(j; J_2; j')$	$\frac{1}{2}\varphi(j_2, j+1)$	$\theta(j_2, j)$	$\frac{1}{2}\varphi(j_2, j)$

TABLE 2
THE MATRIX $(st | s' t')$

$s' \backslash t'$	$t+1$	t	$t-1$
$s+1$	$\frac{1}{2}\xi(s+1, t+1)$	$\eta(s+1, t)/2t(t+1)$	$\frac{1}{2}\zeta(s+1, t-1)$
s	$\frac{1}{2}\varphi(s, t+1)$	$\theta(s, t)$	$\frac{1}{2}\varphi(s, t)$
$s-1$	$\frac{1}{2}\zeta(s, t)$	$\eta(s, t)/2t(t+1)$	$\frac{1}{2}\xi(s, t)$

possible matrix elements will be written down explicitly. In Table 2 the new functions are :

$$\left. \begin{aligned} \xi(s, t) &= \frac{\sqrt{\{P(s, t)P(s, t-1)\}}}{t\sqrt{\{(2t-1)(2t+1)\}}}, \\ \eta(s, t) &= \sqrt{\{P(s, t)Q(s, t)\}}, \\ \zeta(s, t) &= -\frac{\sqrt{\{Q(s, t)Q(s, t+1)\}}}{(t+1)\sqrt{\{(2t+1)(2t+3)\}}} \end{aligned} \right\} \dots\dots (13)$$

(b) *Matrix Elements of \mathbf{J}_k where $\sum_k \mathbf{J}_k = \mathbf{J}$*

We now consider the addition of $n+1$ commuting angular momenta :

$$\sum_{k=0}^n \mathbf{J}_k = \mathbf{J}, \quad \dots\dots\dots (14)$$

and let us add them together one at a time, defining $n-1$ "intermediate" quantum numbers according to

$$\left. \begin{aligned} \mathbf{J}_n + \mathbf{J}_{n-1} &= \mathbf{J}_{n-1, n}, \\ &\dots\dots\dots \\ \mathbf{J}_{k+1, n} + \mathbf{J}_k &= \mathbf{J}_{k, n}, \\ &\dots\dots\dots \\ \mathbf{J}_{1, n} + \mathbf{J}_0 &= \mathbf{J}, \end{aligned} \right\} \dots\dots\dots (15)$$

where \mathbf{J} can be considered to be short for $\mathbf{J}_{0,n}$. Then, since all the \mathbf{J}_k commute, we can put into equation (12)

$$\mathbf{J}_1 \mathbf{J}_2 \mathbf{J} \rightarrow \mathbf{J}_k \mathbf{J}_{k+1, n} \mathbf{J}_{k, n} \mathbf{J}_l \quad (l > k), \quad \dots \dots \dots (16)$$

and it follows that the matrix elements of the components of \mathbf{J}_k are given by the equations (6) together with

$$\begin{aligned} & (j_k, n, j_{k-1}, n \dots j_1, n; j_k, j'_k, n, j_{k-1}, n \dots j_1, n, j') \\ &= (j_k, n; j_k, j'_k, n) (j_k, n, j_{k-1}, n | j'_k, n, j'_{k-1}, n) \dots (j_1, n, j | j'_1, n, j'), \dots \end{aligned} \quad (17)$$

where diagonal quantum numbers (such as $j_{k+1}, j_k, j_{k+1}, n, \dots$) have been omitted, and the elements $(j_k, n, j_{k-1}, n | j'_k, n, j'_{k-1}, n)$ are given in Table 2.

The factors $(j_k, n; j_k, j'_k, n)$ are obtained from Table 1 on putting

$$\mathbf{J}_1 \mathbf{J}_2 \mathbf{J} \rightarrow \mathbf{J}_k \mathbf{J}_{k+1, n} \mathbf{J}_{k, n} \quad (k \neq n), \quad \dots \dots \dots (18)$$

which is consistent with equation (16), and are :

$$\left. \begin{aligned} (j_k, n; j_k, j_{k, n} + 1) &= -\frac{1}{2} \varphi(j_k, j_{k, n} + 1), \\ (j_k, n; j_k, j_{k, n}) &= \theta(j_k, j_{k, n}), \\ (j_k, n; j_k, j_{k, n} - 1) &= -\frac{1}{2} \varphi(j_k, j_{k, n}), \end{aligned} \right\} \quad (k \neq n) \quad \dots \dots (19)$$

while for $k=n$,

$$\left. \begin{aligned} (j_{n-1}, n; j_n, j_{n-1, n} + 1) &= \frac{1}{2} \varphi(j_n, j_{n-1, n} + 1), \\ (j_{n-1}, n; j_n, j_{n-1, n}) &= \theta(j_n, j_{n-1, n}), \\ (j_{n-1}, n; j_n, j_{n-1, n} - 1) &= \frac{1}{2} \varphi(j_n, j_{n-1, n}). \end{aligned} \right\} \quad (k=n) \quad \dots \dots (20)$$

(c) Matrix Elements of $2\mathbf{J}_{k, Z}^2 - \mathbf{J}_{k, X}^2 - \mathbf{J}_{k, Y}^2$ etc.

If we use the equations (6) to express the dependence on m , it turns out that the matrices of $2\mathbf{J}_{k, Z}^2 - \mathbf{J}_{k, X}^2 - \mathbf{J}_{k, Y}^2$ etc. can also be factored into submatrices :

$$\left. \begin{aligned} (\alpha j m | 2\mathbf{J}_{k, Z}^2 - \mathbf{J}_{k, X}^2 - \mathbf{J}_{k, Y}^2 | \alpha' j' m') &= (\alpha j | \mathbf{J}_k^2 | \alpha' j') (j m | j' m') \delta_{m' m}, \\ (\alpha j m | \mathbf{J}_{k, X}^2 - \mathbf{J}_{k, Y}^2 | \alpha' j' m') &= (\alpha j | \mathbf{J}_k^2 | \alpha' j') (j m | j' m') \delta_{m' m \pm 2}, \\ (\alpha j m | \mathbf{J}_{k, X} \mathbf{J}_{k, Y} + \mathbf{J}_{k, Y} \mathbf{J}_{k, X} | \alpha' j' m') &= \pm i (\alpha j | \mathbf{J}_k^2 | \alpha' j') (j m | j' m') \delta_{m' m \pm 2}, \\ (\alpha j m | \mathbf{J}_{k, X} \mathbf{J}_{k, Z} + \mathbf{J}_{k, Z} \mathbf{J}_{k, X} | \alpha' j' m') &= (\alpha j | \mathbf{J}_k^2 | \alpha' j') (j m | j' m') \delta_{m' m \pm 1}, \\ (\alpha j m | \mathbf{J}_{k, Y} \mathbf{J}_{k, Z} + \mathbf{J}_{k, Z} \mathbf{J}_{k, Y} | \alpha' j' m') &= \pm i (\alpha j | \mathbf{J}_k^2 | \alpha' j') (j m | j' m') \delta_{m' m \pm 1}, \end{aligned} \right\} \dots \dots \dots (21)$$

where $j' = j, j \pm 1, j \pm 2$, $m' = m, m \pm 1, m \pm 2$, and $(j m | j' m')$ is given in Table 3, from which the elements $(j m | j + 1 m')$ and $(j m | j + 2 m')$ may be obtained by use of the Hermitian property of the matrix.

In conventional notation the matrix elements of the product of two operators are given by

$$(\alpha | T U | \alpha') = \sum_{\alpha''} (\alpha | T | \alpha'') (\alpha'' | U | \alpha'). \quad \dots \dots \dots (22)$$

and

$$\left. \begin{aligned} (j_{n-1, n} \parallel J_n^2 \parallel j_{n-1, n}) &= \rho(j_n, j_{n-1, n}), \\ (j_{n-1, n} \parallel J_n^2 \parallel j_{n-1, n-1}) &= \frac{1}{2}\chi(j_n, j_{n-1, n}), \quad (k=n) \\ (j_{n-1, n} \parallel J_n^2 \parallel j_{n-1, n-2}) &= \frac{1}{2}\psi(j_n, j_{n-1, n}). \end{aligned} \right\} \dots (27)$$

The simplification of some of the algebra occurring in the calculations leading to equation (24) is assisted by use of the identity (11) together with :

$$\left. \begin{aligned} P(s+q, t) &= P(s, t+q), \\ Q(s+q, t) &= Q(s, t-q), \\ \rho(s, t) &= \theta^2(s, t) - \frac{1}{4}\varphi^2(s, t+1) - \frac{1}{4}\varphi^2(s, t) \\ &= -\gamma^2(s+1, t)/4t^2(t+1)^2 + \frac{1}{4}\xi^2(s+1, t+1) + \frac{1}{4}\zeta^2(s+1, t-1) \\ &= -\gamma^2(s, t)/4t^2(t+1)^2 + \frac{1}{4}\xi^2(s, t) + \frac{1}{4}\zeta^2(s, t). \end{aligned} \right\} \dots (28)$$

TABLE 4
THE MATRIX ($s \parallel s' t'$)

$s' \backslash s$	t	$t-1$	$t-2$
$s+2$	$3\eta(s+1, t)\eta(s+2, t)\Phi(t)$	$\frac{1}{2}\zeta(s+2, t-1)\eta(s+1, t)\Phi(t-1)$	$\frac{1}{2}\zeta(s+1, t-1)\zeta(s+2, t-2)$
$s+1$	$3\eta(s+1, t)[R(s, t)+s]\Phi(t)$	$\frac{1}{2}\zeta(s+1, t-1)[R(s, t)+s(t+1)]\Phi(t-1)$	$\frac{1}{2}\varphi(s, t)\zeta(s+1, t-2)$
s	$\rho(s, t)$	$\frac{1}{2}\chi(s, t)$	$\frac{1}{2}\psi(s, t)$
$s-1$	$3\eta(s, t)[R(s-1, t)+(s-1)]\Phi(t)$	$\frac{1}{2}\xi(s, t)[R(s, t)-(s+1)(t+1)]\Phi(t-1)$	$\frac{1}{2}\varphi(s, t)\xi(s, t-1)$
$s-2$	$3\eta(s, t)\eta(s-1, t)\Phi(t)$	$\frac{1}{2}\xi(s, t)\eta(s-1, t-1)\Phi(t-1)$	$\frac{1}{2}\xi(s, t)\xi(s-1, t-1)$

(d) *Matrix Elements of $2J_{k, Z}J_Z - J_{k, X}J_X - J_{k, Y}J_Y$ etc.*

Using the matrix elements (4) and (17), it follows easily that

$$\begin{aligned} & (j_k, n j_{k-1, n} \dots j_1, n j_m \mid 2J_{k, Z}J_Z - J_{k, X}J_X - J_{k, Y}J_Y \mid j'_k, n j'_{k-1, n} \dots j'_1, n j' m') \\ &= (j_k, n j_{k-1, n} \dots j_1, n j \mid J_k \mid j'_k, n j'_{k-1, n} \dots j'_1, n j') (j_m \parallel j' m') \delta_{m'm} \\ &= (j_k, n \mid J_k \mid j'_k, n) (j_k, n j_{k-1, n} \mid j'_k, n j'_{k-1, n}) \dots (j_1, n j \mid j'_1, n j') (j_m \parallel j' m') \delta_{m'm}, \\ & \dots (29) \end{aligned}$$

with the factors already defined, and with $j'_{k, n} = j_{k, n}, j_{k, n} \pm 1$, etc.

The other functions of the components of J_k have matrix elements differing only in their dependence on j and m , as in equation (21).

(e) *Matrix Elements of $2J_{k, Z}J_{l, Z} - J_{k, X}J_{l, X} - J_{k, Y}J_{l, Y}$ etc. ($l > k$)*

These cross-product terms have matrix elements of the form

$$\begin{aligned} & (j_l, n \dots j_k, n \dots j_m \mid 2J_{k, Z}J_{l, Z} - J_{k, X}J_{l, X} - J_{k, Y}J_{l, Y} \mid j'_l, n \dots j'_k, n \dots j' m') \\ &= (j_l, n \dots j_k, n \dots j \parallel J_k J_l \parallel j'_l, n \dots j'_k, n \dots j') (j_m \parallel j' m') \delta_{m'm} \\ &= (j_l, n \mid J_l \mid j'_l, n) (j_l, n j_{l-1, n} \mid j'_l, n j'_{l-1, n}) \dots (j_k, n \mid J_k \mid j'_k, n) \times \\ & \quad (j_k, n j_{k-1, n} \mid j'_k, n j'_{k-1, n}) \dots (j_1, n j \mid j'_1, n j') (j_m \parallel j' m') \delta_{m'm} \dots (30) \end{aligned}$$

In the particular case $k=n-1$ and $l=n$ we get

$$\begin{aligned} & \langle j_{n-1}, n \dots j_1, n | j m | 2J_{n-1, Z} J_n, Z - J_{n-1, X} J_n, X - J_{n-1, Y} J_n, Y | j'_{n-1}, n \dots j'_1, n | j' m' \rangle \\ &= (j_{n-1}, n || J_{n-1} J_n || j_{n-1}, n-1) (j_{n-1}, n | j_{n-2}, n || j'_{n-1}, n | j'_{n-2}, n) \dots \times \\ & (j_1, n | j'_1, n) (j m || j' m') \delta_{m' m}, \dots \dots \dots (31) \end{aligned}$$

with

$$\left. \begin{aligned} (j_{n-1}, n || J_{n-1} J_n || j_{n-1}, n) &= \theta(j_{n-1}, j_{n-1}, n) - \rho(j_{n-1}, j_{n-1}, n), \\ (j_{n-1}, n || J_{n-1} J_n || j_{n-1}, n-1) &= -\frac{1}{2} \varphi(j_{n-1}, j_{n-1}, n) + \frac{1}{2} \chi(j_{n-1}, j_{n-1}, n), \\ (j_{n-1}, n || J_{n-1} J_n || j_{n-1}, n-2) &= -\frac{1}{2} \psi(j_{n-1}, j_{n-1}, n). \end{aligned} \right\} \dots \dots \dots (32)$$

When $l=k+1 \neq n$ there also results

$$\begin{aligned} & \langle j_{k+1}, n \dots j m | 2J_{k, Z} J_{k+1, Z} - J_{k, X} J_{k+1, X} - J_{k, Y} J_{k+1, Y} | j'_{k+1}, n \dots j' m' \rangle \\ &= (j_{k+1}, n | j_{k+1}, n) [j_{k+1}, n | j'_k, n || j'_{k+1}, n | j'_k, n] \times \\ & (j_k, n | j_{k-1}, n || j'_k, n | j'_{k-1}, n) \dots (j_1, n | j'_1, n) (j m || j' m') \delta_{m' m}, \dots (33) \end{aligned}$$

with $[j_{k+1}, n | j_k, n || j'_{k+1}, n | j'_k, n]$ given in Table 5.

TABLE 5
THE MATRIX $[j_{k+1}, n | j_k, n || j'_{k+1}, n | j'_k, n]$

$j'_{k+1}, n \backslash j'_k, n$	j_k, n	$j_k, n-1$	$j_k, n-2$
$j_{k+1}, n+1$	$-\eta(j_{k+1}, n+1, j_k, n) \times$ $[3R(j_{k+1}, n, j_k, n)$ $+ 3(j_{k+1}, n+1)$ $- 4j_k, n(j_k, n+1)] \times$ $\Phi(j_k, n)$	$\frac{1}{2} \zeta(j_{k+1}, n+1, j_k, n-1) \times$ $\{1 - [R(j_{k+1}, n, j_k, n)$ $+ j_{k+1}, n(j_k, n+1)] \times$ $\Phi(j_k, n-1)\}$	$-\frac{1}{2} \zeta(j_{k+1}, n+1, j_k, n-2) \times$ $\varphi(j_{k+1}, n, j_k, n)$
j_{k+1}, n	$\theta(j_{k+1}, n, j_k, n)$ $-\rho(j_{k+1}, n, j_k, n)$	$\frac{1}{2} \varphi(j_{k+1}, n, j_k, n)$ $-\frac{1}{2} \chi(j_{k+1}, n, j_k, n)$	$-\frac{1}{2} \psi(j_{k+1}, n, j_k, n)$
$j_{k+1}, n-1$	$-\eta(j_{k+1}, n, j_k, n) \times$ $[3R(j_{k+1}, n-1, j_k, n)$ $+ 3j_{k+1}, n$ $- 4j_k, n(j_k, n+1)] \times$ $\Phi(j_k, n)$	$\frac{1}{2} \xi(j_{k+1}, n, j_k, n) \times$ $\{1 - [R(j_{k+1}, n, j_k, n)$ $-(j_{k+1}, n+1)(j_k, n+1)]$ $\times \Phi(j_k, n-1)\}$	$-\frac{1}{2} \xi(j_{k+1}, n, j_k, n) \times$ $\varphi(j_{k+1}, n-1, j_k, n-1)$

(f) *Matrix Elements of* $J_{k, X} J_X + J_{k, Y} J_Y + J_{k, Z} J_Z$

This function is the scalar product $J_k \cdot J$, and its matrix elements are diagonal in j (TAS Sec. 8³). Also,

$$J_k \cdot J = 3J_{k, Z} J_Z - (2J_{k, Z} J_Z - J_{k, X} J_X - J_{k, Y} J_Y),$$

and it follows from previous results that

$$\begin{aligned} & \langle j_k, n \dots j_1, n | j m | J_{k, X} J_X + J_{k, Y} J_Y + J_{k, Z} J_Z | j'_k, n \dots j'_1, n | j' m' \rangle \\ &= (j_k, n | j'_k, n) j_k, n | j_{k-1}, n || j'_k, n | j'_{k-1}, n) \dots (j_1, n | j'_1, n) \times j(j+1) \delta_{j' j} \delta_{m' m}. \end{aligned} \dots \dots \dots (34)$$

IV. MATRIX ELEMENTS FOR ONE NUCLEAR SPIN

When only one nuclear spin $\mathbf{I}=\mathbf{I}_1$ is present in the molecule we have the coupling scheme

$$\mathbf{J}+\mathbf{I}=\mathbf{F}, \quad \dots\dots\dots (35)$$

or, using the reversed spin angular momentum $\bar{\mathbf{I}}$,

$$\mathbf{F}+\bar{\mathbf{I}}=\mathbf{J}, \quad \dots\dots\dots (36)$$

in which the components of \mathbf{F} and $\bar{\mathbf{I}}$ commute (Van Vleck 1951).

The Hamiltonian (2) is referred to the molecular axis system, and thus (as shown, for example, by Van Vleck (1951)) the results of Section III apply if we change the sign of i and replace m by the quantum number K . Thus our representation may be labelled by the quantum numbers IJK .

We can now use the results of Section III, putting $n=1$ and (cf. equation (18))

$$\mathbf{J}_k \mathbf{J}_{k+1}, n \mathbf{J}_{k,n} \rightarrow \bar{\mathbf{I}} \mathbf{F} \mathbf{J}, \quad \dots\dots\dots (37)$$

with $k=0$. Then the matrix elements of the Hamiltonian (2) can be written

$$(JK | H | J' K') = \{(K | 1 | K') (J || \bar{I}^2 || J') + (K | 1' | K') (J || \bar{I} || J')\} (JK || J' K'),$$

..... (38)

where $K'=K, K\pm 1, K\pm 2$, and $J'=J, J\pm 1, J\pm 2$. We have introduced the symbols 1 and $1'$ in equation (38) to distinguish between the various kinds of terms contributing to the Hamiltonian. $(JK || J' K')$ is given by Table 6,

TABLE 6
THE MATRIX $(JK || J' K')$

$J' \backslash K'$	K	$K\pm 1$	$K\pm 2$
J	$3K^2 - J(J+1)$	$\frac{1}{2}(2K\pm 1)f(J, \pm K)$	$\frac{1}{2}f(J, \pm K)f(J, \pm K+1)$
$J-1$	$3K\sqrt{(J^2-K^2)}$	$\frac{1}{2}(J\pm 2K+1)g(J, \pm K)$	$\pm \frac{1}{2}f(J, \pm K)g(J-1, \pm K)$
$J-2$	$3g(J, K)g(J, -K)$	$\pm g(J, \pm K+1)\sqrt{(J^2-K^2)}$	$\frac{1}{2}g(J, \pm K)g(J, \pm K+2)$

which is just Table 3 with j and m replaced by J and K . The elements of Table 6 agree with the phase convention of TAS; if it is required that the matrix elements of this paper be consistent with those of Cross, Hainer, and King (1944), then the column $K\pm 1$ would have to be multiplied by $\pm i$, and the column $K\pm 2$ by -1 . $(J || \bar{I}^2 || J')$ and $(J || \bar{I} || J')$ come from equations (26) and (19) respectively, with $j_k j_{k,n} \rightarrow IJ$, and we get the results of Tables 7 and 8. The quantities in Table 7 are the coefficients of the operators in the Hamiltonian, and, since there is only one spin, the subscripts etc. have been dropped. These coefficients are defined

in equation (3). The functions occurring in Table 8 have been defined in equations (9) and (25); explicitly, they are:

$$\left. \begin{aligned} \rho(I, J) &= \{3R(I, J)[R(I, J) - 1] - 4I(I+1)J(J+1)\}\Phi(J), \\ \chi(I, J) &= [\theta(I, J) + \theta(I, J-1)]\varphi(I, J), \\ \psi(I, J) &= \frac{1}{2}\varphi(I, J)\varphi(I, J-1), \\ \Phi(J) &= [2J(J+1)(2J-1)(2J+3)]^{-1} \\ \varphi(I, J) &= -\frac{\sqrt{\{P(I, J)Q(I, J-1)\}}}{J\sqrt{\{(2J-1)(2J+1)\}}}, \\ \theta(I, J) &= \frac{R(I, J)}{2J(J+1)}, \end{aligned} \right\} \dots (39)$$

with

$$\left. \begin{aligned} P(I, J) &= (J-F+I)(J+F+I+1), \\ Q(I, J) &= (F+I-J)(J+F-I+1), \\ R(I, J) &= J(J+1) - F(F+1) + I(I+1). \end{aligned} \right\} \dots \dots \dots (40)$$

The above results have already been given by Van Vleck (1951) except that here we have a negative sign in the $(J | J \pm 1)$ matrix elements; this arises

TABLE 7
THE FACTORS $(K | I | K')$ AND $(K | I' | K')$

K'	K	$K \pm 1$	$K \pm 2$
$(K I K')$	α	$\delta \mp i\epsilon$	$\beta \mp i\gamma$
$(K I' K')$	$\alpha + \frac{J(J+1)\alpha'}{3K^2 - J(J+1)}\delta_{J'J}$	$d \mp ie$	$b \mp ic$

TABLE 8
THE MATRICES $(J || \hat{I}^2 || J')$ AND $(J; \hat{I}; J')$

J'	J	$J-1$	$J-2$
$(J \hat{I}^2 J')$	$\rho(I, J)$	$-\frac{1}{2}\chi(I, J)$	$\frac{1}{2}\psi(I, J)$
$(J; \hat{I}; J')$	$\theta(I, J)$	$-\frac{1}{2}\varphi(I, J)$	0

from our association of I with j_1 rather than with j_2 as in Van Vleck's paper, for we have had to be consistent in our notation (cf equations (16) and (18)) in order to deal with the cross-product terms (Section III (e)) correctly.

It will be observed that

$$R(I, J) = 2\bar{\mathbf{I}} \cdot \mathbf{J} = -2\mathbf{I} \cdot \mathbf{J} = -C, \dots \dots \dots (41)$$

where C is common notation. Then,

$$\theta(I, J) = \frac{2\bar{\mathbf{I}} \cdot \mathbf{J}}{2J(J+1)}, \dots \dots \dots (42)$$

Also,

$$\rho(I, J) = \frac{4I(2I-1)}{J(J+1)} f(I, J, F), \quad \dots \quad (43)$$

where $f(I, J, F)$, sometimes called "Casimir's function", is tabulated by Townes and Schawlow (1955, Appendix I).

For symmetric top molecules the only non-vanishing matrix elements are those diagonal in K , and are:

$$\left. \begin{aligned} (JK | H | JK) &= [3K^2 - J(J+1)] \alpha \rho(I, J) \\ &\quad + \{a'J(J+1) + a[3K^2 - J(J+1)]\} \theta(I, J), \\ (JK | H | J-1K) &= -\frac{3}{2}K \sqrt{J^2 - K^2} [\alpha \chi(I, J) + a \varphi(I, J)], \\ (JK | H | J-2K) &= \frac{3}{2}g(J, K)g(J, -K) \alpha \psi(I, J). \end{aligned} \right\} \dots \quad (44)$$

In an asymmetric top molecule if only a' , a , b , α , and β need be considered (because of molecular symmetry or otherwise), the first-order matrix elements diagonal in J and τ become (Bragg 1948):

$$(J\tau | H | J\tau) = \alpha_{J,\tau} \rho(I, J) + c_{J,\tau} \theta(I, J), \quad \dots \quad (45)$$

where

$$\left. \begin{aligned} \alpha_{J,\tau} &= \sum_g \chi_{gg} \langle J_g^2 \rangle / 2I(2I-1), \\ c_{J,\tau} &= \sum_g M_{gg} \langle J_g^2 \rangle, \\ \chi_{gg} &= eQ \langle \partial^2 V / \partial g^2 \rangle, \\ \langle J_a^2 \rangle &= \frac{1}{2} [J(J+1) + E - (\kappa+1) \partial E / \partial \kappa], \\ \langle J_b^2 \rangle &= \partial E / \partial \kappa, \\ \langle J_c^2 \rangle &= \frac{1}{2} [J(J+1) - E + (\kappa-1) \partial E / \partial \kappa], \end{aligned} \right\} \dots \quad (46)$$

in which a , b , and c refer to the principal inertial axes, E is the reduced rigid rotor energy, and κ is the asymmetry parameter.

Usually Laplace's equation $\sum_g \partial^2 V / \partial g^2 = 0$ is applied to eliminate one of the χ_{gg} .

V. MATRIX ELEMENTS FOR TWO NUCLEAR SPINS

(a) Matrix Elements for Similar Coupling: $I_2 I_1 I J F K$ Representation

When the energy of interaction with the rest of the molecule is about the same for each nucleus, the most appropriate coupling scheme is (Foley 1947)

$$\left. \begin{aligned} I_2 + I_1 &= I, \\ I + J &= F, \end{aligned} \right\} \dots \quad (47)$$

where I is the total spin angular momentum. The intermediate and total angular momentum quantum numbers cover the ranges

$$\left. \begin{aligned} I &= I_1 + I_2, I_1 + I_2 - 1, \dots, |I_1 - I_2|, \\ F &= J + I, J + I - 1, \dots, |J - I|. \end{aligned} \right\} \dots \quad (48)$$

In general I is not a good quantum number.

By using the reversed spin angular momenta to give

$$\left. \begin{aligned} \bar{\mathbf{I}}_2 + \bar{\mathbf{I}}_1 &= \bar{\mathbf{I}}, \\ \bar{\mathbf{I}} + \mathbf{F} &= \mathbf{J}, \end{aligned} \right\} \dots\dots\dots (49)$$

in the molecular representation $I_2 I_1 I J F K$, and putting $n=2$ and

$$\mathbf{J}_2 \mathbf{J}_1 \mathbf{J}_{1,2} \mathbf{J}_0 \mathbf{J} \rightarrow \bar{\mathbf{I}}_2 \bar{\mathbf{I}}_1 \bar{\mathbf{I}} \mathbf{F} \mathbf{J} \dots\dots\dots (50)$$

in equation (16), we can write, similarly to equation (38),

$$\begin{aligned} (IJK | H | I' J' K') &= [\Sigma_{j=1,2} s(K | l_j | K') (I || \bar{I}_j^2 || I') (IJ || I' J') \\ &\quad + \Sigma_{j=1,2} (K | l_j | K') (I : \bar{I}_j : I') (IJ | I' J')] (JK || J' K'), \end{aligned} \dots\dots\dots (51)$$

in which $I'=I, I\pm 1, I\pm 2$, $J'=J, J\pm 1, J\pm 2$, and $K'=K, K\pm 1, K\pm 2$. The matrix $(JK || J' K')$ has been given in Table 6, and $(K | l_j | K')$ and $(K | l_j' | K')$

TABLE 9
THE FACTORS $(K | l_j | K')$ AND $(K | l_j' | K')$

K'	K	$K\pm 1$	$K\pm 2$
$(K l_1 K')$	α_1	$\delta_1 \mp i\epsilon_1$	$\beta_1 \mp i\gamma_1$
$(K l_2 K')$	α_2	$\delta_2 \mp i\epsilon_2$	$\beta_2 \mp i\gamma_2$
$(K l_S K')$	α_S	$\delta_S \mp i\epsilon_S$	$\beta_S \mp i\gamma_S$
$(K l_1' K')$	$a_1 + \frac{J(J+1)\alpha_1'}{3K^2 - J(J+1)} \delta_{J'J}$	$d_1 \mp i\epsilon_1$	$b_1 \mp i\epsilon_1$
$(K l_2' K')$	$a_2 + \frac{J(J+1)\alpha_2'}{3K^2 - J(J+1)} \delta_{J'J}$	$d_2 \mp i\epsilon_2$	$b_2 \mp i\epsilon_2$

TABLE 10
THE MATRICES $(I || \bar{I}_j^2 || I')$ AND $(I : \bar{I}_j : I')$

I'	I	$I-1$	$I-2$
$(I \bar{I}_1^2 I')$	$\rho(I_1, I)$	$-\frac{1}{2}\chi(I_1, I)$	$\frac{1}{2}\psi(I_1, I)$
$(I \bar{I}_2^2 I')$	$\rho(I_2, I)$	$\frac{1}{2}\chi(I_2, I)$	$\frac{1}{2}\psi(I_2, I)$
$(I \bar{I}_S^2 I')$	$\theta(I_1, I) - \rho(I_1, I)$	$-\frac{1}{2}\varphi(I_1, I) + \frac{1}{2}\chi(I_1, I)$	$-\frac{1}{2}\psi(I_1, I)$
$(I : \bar{I}_1 : I')$	$\theta(I_1, I)$	$-\frac{1}{2}\varphi(I_1, I)$	0
$(I : \bar{I}_2 : I')$	$\theta(I_2, I)$	$\frac{1}{2}\varphi(I_2, I)$	0

are shown in Table 9, where we have omitted a superscript on the α_S etc. since there is only one dipole-dipole term. The factors $(I || \bar{I}_j^2 || I')$ (with $\bar{I}_S^2 \equiv \bar{I}_1 \bar{I}_2$) and $(I : \bar{I}_j : I')$ are obtained from equations (26), (27), (32), (19), and (20), using (50), and are shown in Table 10.

It should be noted that, from equations (11) and (25),

$$\psi(I_2, I) = \psi(I_1, I). \quad \dots\dots\dots (52)$$

$(IJ \parallel I' J')$ and $(IJ \mid I' J')$ are obtained by putting $st \rightarrow IJ$ in Tables 4 and 2 respectively, and need not be rewritten here.

For an asymmetric top molecule the first-order matrix elements diagonal in J and τ become (cf. equation (45)):

$$\left. \begin{aligned} (IJ\tau \mid H \mid IJ\tau) &= \{ \alpha_{J,\tau}^{(1)} \rho(I_1, I) + \alpha_{J,\tau}^{(2)} \rho(I_2, I) + \alpha_{J,\tau}^{(S)} [\theta(I_1, I) - \rho(I_1, I)] \} \rho(I, J) \\ &\quad + [c_{J,\tau}^{(1)} \theta(I_1, I) + c_{J,\tau}^{(2)} \theta(I_2, I)] \theta(I, J), \\ (IJ\tau \mid H \mid I-1J\tau) &= -\{ \alpha_{J,\tau}^{(1)} \chi(I_1, I) - \alpha_{J,\tau}^{(2)} \chi(I_2, I) \\ &\quad + \alpha_{J,\tau}^{(S)} [\varphi(I_1, I) - \chi(I_1, I)] \} \frac{3}{2} \eta(I, J) [R(I, J) - (I+1)] \Phi(J) \\ &\quad - (c_{J,\tau}^{(1)} - c_{J,\tau}^{(2)}) \varphi(I_1, I) \eta(I, J) / 4J(J+1), \\ (IJ\tau \mid H \mid I-2J\tau) &= (\alpha_{J,\tau}^{(1)} + \alpha_{J,\tau}^{(2)} - \alpha_{J,\tau}^{(S)}) \frac{3}{2} \psi(I_1, I) \eta(I, J) \eta(I-1, J) \Phi(J), \\ &\dots\dots\dots \end{aligned} \right\} \quad (53)$$

where

$$\left. \begin{aligned} \alpha_{J,\tau}^{(k)} &= \sum_g \chi_{gg}^{(k)} \langle J_g^2 \rangle / 2I_k(2I_k-1), \\ \alpha_{J,\tau}^{(S)} &= -(3\mu_N^2 g_1 g_2) [\sum_g \langle r_g^2 / r^5 \rangle \langle J_g^2 \rangle - \frac{1}{3} \langle r^{-3} \rangle J(J+1)], \\ c_{J,\tau}^{(k)} &= \sum_g M_{gg}^{(k)} \langle J_g^2 \rangle, \\ \chi_{gg}^{(k)} &= e Q_k \langle \partial^2 V_k / \partial g^2 \rangle, \end{aligned} \right\} \quad \dots (54)$$

the $\langle J_g^2 \rangle$ being defined in equation (46) and r standing for r_{12} .

When $I_1 = I_2$ the quadrupolar parts of equation (53) reduce to the matrix elements of the special cases discussed by Robinson and Cornwell (1953) and by Ramsey (1956, pp. 63, 83).

If the two spins are equal and have equal couplings α_k , the matrix elements of the quadrupole coupling off-diagonal by 1 in I will vanish. If also the spins are symmetrically placed in the molecule, as, for example, in C_{2v} molecules, energy levels with total spin I differing by one unit cannot exist simultaneously for a given symmetry of J and τ , hence the off-diagonal matrix elements $(I \mid I \pm 1)$ can connect only states which also differ in J and/or K . In general, then, if this spin symmetry exists, these $(I \mid I \pm 1)$ contributions to the hyperfine energies will be off-diagonal in the rotation energy and may usually be neglected. However, in a symmetric top molecule levels of different K are degenerate, and the coupling between such states will be first-order, so the splitting may be appreciable, as in the case of ammonia (Gunther-Mohr, Townes, and Van Vleck 1954).

The relative intensities of transitions between hyperfine components described by this coupling scheme can be computed by the method described by Robinson and Cornwell (1953).

The effects of matrix elements off-diagonal in J may be calculated by perturbation theory; in many cases the comparatively large spacings of rotational energies will enable these effects to be neglected.

To first order, only the diagonal elements of the tensor $\chi_{qq'}^{(k)}$ contribute to the energy (Bragg 1948). If the principal axes of this tensor do not coincide with the inertial principal axes x , y , and z , then, in higher order, the off-diagonal elements may have to be taken into account.

(b) *Matrix Elements for Non-similar Coupling: $I_1 J F_1 I_2 F_2 K$ Representation*

When one spin (I_1) is more strongly coupled to the rotation than is the other it is usually more convenient to use the coupling scheme (Bardeen and Townes 1948a)

$$\left. \begin{array}{l} \mathbf{J} + \mathbf{I}_1 = \mathbf{F}_1, \\ \mathbf{F}_1 + \mathbf{I}_2 = \mathbf{F}, \end{array} \right\} \dots\dots\dots (55)$$

with

$$\left. \begin{array}{l} F_1 = J + I_1, \quad J + I_1 - 1, \dots, |J - I_1|, \\ F = F_1 + I_2, \quad F_1 + I_2 - 1, \dots, |F_1 - I_2|. \end{array} \right\} \dots\dots (56)$$

In such a representation the general matrix elements of the Hamiltonian become rather more complicated than the previous example, because of the lack of symmetry between I_1 and I_2 :

$$\begin{aligned} (F_1 J K | H | F_1' J' K') = & \\ & \{ (K | I_1 | K') (J || \bar{I}_1^2 || J') + (K | I_2 | K') (F_1 || \bar{I}_2^2 || F_1') (F_1 J || F_1' J') \\ & + (K | I_1 | K') (F_1 || \bar{I}_2^2 || F_1') [F_1 J || F_1' J'] + (K | I_1 | K') (J || \bar{I}_1^2 || J') \\ & + (K | I_2 | K') (F_1 || \bar{I}_2^2 || F_1') (F_1 J || F_1' J') \} (J K || J' K'), \dots\dots\dots (57) \end{aligned}$$

with $(K | I_j | K')$, $(K | I_j^2 | K')$, and $(J K || J' K')$ already given in Tables 9 and 6, respectively, $(J || \bar{I}_1^2 || J')$ and $(F_1 || \bar{I}_2^2 || F_1')$ given in equation (26), $(J || \bar{I}_1^2 || J')$ and $(F_1 || \bar{I}_2^2 || F_1')$ given by equation (19), and $(F_1 J || F_1' J')$, $(F_1 J || F_1' J')$, and $[F_1 J || F_1' J']$ obtained from Tables 2, 4, and 5 respectively.

The matrix elements diagonal in J and τ are (cf. equation (53)):

$$\begin{aligned} (F_1 J \tau | H | F_1 J \tau) = & \alpha_{J, \tau}^{(1)} \rho(I_1, J) + \alpha_{J, \tau}^{(2)} \rho(I_2, F_1) \rho(F_1, J) \\ & + \alpha_{J, \tau}^{(S)} \theta(I_2, F_1) [\theta(F_1, J) - \rho(F_1, J)] \\ & + c_{J, \tau}^{(1)} \theta(I_1, J) + c_{J, \tau}^{(2)} \theta(I_2, F_1) \theta(F_1, J), \\ (F_1 J \tau | H | F_1 - 1 J \tau) = & -\frac{3}{2} \alpha_{J, \tau}^{(2)} \chi(I_2, F_1) \eta(F_1, J) [R(F_1 - 1, J) + (F_1 - 1)] \Phi(J) \\ & + \frac{1}{2} \alpha_{J, \tau}^{(S)} \varphi(I_2, F_1) \eta(F_1, J) [3R(F_1 - 1, J) + 3F_1 - 4J(J + 1)] \Phi(J) \\ & - c_{J, \tau}^{(2)} \varphi(I_2, F_1) \eta(F_1, J) / 4J(J + 1), \\ (F_1 J \tau | H | F_1 - 2 J \tau) = & \frac{3}{2} \alpha_{J, \tau}^{(2)} \psi(I_2, F_1) \eta(F_1, J) \eta(F_1 - 1, J) \Phi(J). \end{aligned} \dots\dots\dots (58)$$

When the coupling of nucleus 2 is much less than that of nucleus 1 it is frequently a good approximation to ignore the matrix elements off-diagonal in F_1 , when also the relative intensities of transitions are readily obtained from tables such as in TAS.

VI. CONCLUSION

It will be apparent that for any number of nuclear spins the matrix elements of the Hamiltonian can be written down from the above results and converted to numerical form for particular values of the quantum numbers. However, it is another matter to determine in advance which of these matrix elements vanish because of the symmetry of the molecule.

The choice of a suitable representation which will simplify diagonalization of the resultant matrices depends on the specific problem, and has been discussed for some examples by Bersohn (1949).

VII. ACKNOWLEDGMENTS

The author would like to thank Mr. W. E. Smith for valuable criticism and suggestions.

VIII. REFERENCES

- BARDEEN, J., and TOWNES, C. H. (1948a).—*Phys. Rev.* **73**: 97–105.
BARDEEN, J., and TOWNES, C. H. (1948b).—*Phys. Rev.* **73**: 627–9, 1204.
BERSOHN, R. (1949).—Ph.D. Thesis, Harvard University.
BERSOHN, R. (1950).—*J. Chem. Phys.* **18**: 1124–5.
BRAGG, J. K. (1948).—*Phys. Rev.* **74**: 533–8.
CONDON, E. U., and SHORTLEY, G. H. (1953).—“The Theory of Atomic Spectra.” (Cambridge Univ. Press.)
CROSS, P. C., HAINER, R. M., and KING, G. W. (1944).—*J. Chem. Phys.* **12**: 210–43.
FOLEY, H. M. (1947).—*Phys. Rev.* **71**: 747–51.
GORDON, J. P. (1955).—*Phys. Rev.* **99**: 1253–63.
GUNTHER-MOHR, G. R., TOWNES, C. H., and VAN VLECK, J. H. (1954).—*Phys. Rev.* **94**: 1191–203.
GÜTTINGER, P., and PAULI, W. (1931).—*Z. Phys.* **67**: 743–65.
HADLEY, G. F. (1955).—Ph.D. Thesis, Harvard University.
HERRMANN, G. (1956).—New York University Rep. No. 289.6.
MYERS, R. J., and GWINN, W. D. (1952).—*J. Chem. Phys.* **20**: 1420–7.
OKAYA, A. (1956).—*J. Phys. Soc. Japan* **11**: 249–58.
RACAH, G. (1942).—*Phys. Rev.* **62**: 438–62.
RAMSEY, N. F. (1956).—“Molecular Beams.” (Clarendon Press: Oxford.)
ROBINSON, G. W., and CORNWELL, C. D. (1953).—*J. Chem. Phys.* **21**: 1436–42.
TOWNES, C. H., and SCHAWLOW, A. L. (1955).—“Microwave Spectroscopy.” (McGraw-Hill: New York.)
VAN VLECK, J. H. (1951).—*Rev. Mod. Phys.* **23**: 213–27.
WHITE, R. L. (1955).—*Rev. Mod. Phys.* **27**: 276–88.

THE TOWNSEND IONIZATION COEFFICIENTS IN CROSSED ELECTRIC AND MAGNETIC FIELDS

By H. A. BLEVIN* and S. C. HAYDON*

[Manuscript received October 22, 1957]

Summary

An expression is obtained for the first Townsend ionization coefficient in uniform crossed electric and magnetic fields, and shown to be in better agreement with observation than previous theoretical expressions. The "equivalent pressure" concept for the effect of a transverse magnetic field on this coefficient is shown to be a valid approach to the problem, although the value for the equivalent pressure obtained in this analysis differs from the values given by earlier authors.

The effect of a transverse magnetic field upon the second Townsend coefficient is discussed in greater detail than hitherto, and the possibility of differentiating by this means between the secondary processes operating is discussed.

I. INTRODUCTION

The mechanism of electrical breakdown of gases in uniform static electric fields is now well established in terms of the growth of current equations based on the first and second Townsend ionization coefficients, α and ω/α . The more complex situation in the presence of crossed electric and magnetic fields is not so well understood, however, and previous attempts (Wehrli 1922 ; Valle 1950 ; Somerville 1952 ; Haefer 1953) to explain the observed breakdown characteristics have not been entirely satisfactory. Each of these investigations has been concerned with the influence of a transverse magnetic field on one or both of the Townsend ionization coefficients.

Wehrli made calculations of the first Townsend coefficient in uniform crossed electric and magnetic fields, basing his theory on the assumption that all electron collisions with gas molecules are completely inelastic and that the free path l is constant for all electrons. In this case, an electron will describe a cycloidal path between collisions.

The distance l' travelled in the direction of the electric field E (V/cm) will then be

$$l' = l(1 - eH^2l/8 \times 10^8 Em), \dots\dots\dots (1.01)$$

H being the magnetic intensity in oersteds, e and m the charge (e.m.u.) and mass (g) of an electron. When the magnetic field is absent, $l' = l$ and, since only the component of the free path in the direction of the field E affects the kinetic energy of the electron, Wehrli concluded that, in this sense, the effect of the magnetic field is equivalent to an increase in the pressure to a value p_e , where

$$p_e = p/(1 - eH^2l/8 \times 10^8 Em). \dots\dots\dots (1.02)$$

* Physics Department, University of New England, Armidale, N.S.W.

Valle inserted this "equivalent pressure" p_e into Townsend's approximate expression for α/p in static electric fields, namely,

$$\alpha/p = A \exp(-Bp/E), \dots\dots\dots (1.03)$$

and so obtained an empirical relationship for α in a magnetic field. Neither of these authors considered the effect of the magnetic field on the secondary coefficient.

Valle's theory has been discussed in detail by Somerville (1952) and Haefer (1953) and they have shown that the neglect of the distribution of free paths about the mean is the cause of many of the major qualitative differences between the theoretical and observed values of sparking potentials in crossed electric and magnetic fields. By considering this distribution of free paths, Somerville obtained an improved form of the equivalent pressure concept, which is in better though not entirely satisfactory agreement with experiment.

As an alternative approach, Somerville (1952) and Haefer (1953) independently derived a new expression for α/p in crossed fields without resorting to the equivalent pressure concept. This derivation was based on the assumption of completely inelastic collisions and an ionization probability of unity for all those collisions for which the electron energy is greater than the ionization energy of the gas molecule. Their expression is

$$\alpha/p = A \sinh \{ (a/2l) \sqrt{1 - 4BL/Ea} \} / \varphi(l/a) \sinh(a/2l), \dots (1.04)$$

where a is the length of a complete cycloidal arch described by an electron starting from rest, i.e. $a = 8 \times 10^8 Em/eH^2$,

L is the mean free path at 1 mm Hg pressure ($L = pl$),

A, B are the empirical constants occurring in equation (1.03), and

$$\varphi(x) = \coth(1/2x) - 2x.$$

Furthermore, Somerville has shown that, provided $4BL/Ea < 1$, this expression for α/p leads to sparking potentials not greatly different from those deduced from his modification to Valle's theory, but that for $4BL/Ea > 1$ the theory breaks down completely, because the maximum energy gained over a cycloidal path by an electron starting from rest is then always less than the ionization potential of the gas, and, with the assumptions made, no ionization can then occur. In this case an adequate theory must take account of the possibility of an electron obtaining sufficient energy to ionize as the result of energy gained over several free paths.

Somerville (1952) and Haefer (1953) further extended their investigations to include the effect of a transverse magnetic field on Townsend's secondary coefficient ω/α , but limited their discussion to the case when positive ion action, either at the cathode or in the gas, is the only secondary process. In this case the problem is greatly simplified, since the magnetic field has little influence on the motion of positive ions in the gas and consequently their energies will not be changed appreciably except at very high magnetic field strengths. For this simplified case, the problem reduces to an investigation of the effect of the magnetic field on the secondary electrons produced at the cathode by the positive

ions. The magnetic field causes these electrons to move in cycloidal paths and they may be recaptured by the cathode, thus effectively reducing ω/α .

All these previous theories have approached the problem through the study of individual electron trajectories in the gas. They have entailed some drastic simplifying assumptions, however, and, while qualitative agreement with experiment is reasonably good, quantitative agreement is far from satisfactory. This is hardly surprising when it is realized that, as yet, a satisfactory theory has not been derived on this basis even in the absence of a magnetic field. In view of this, a new approach is made to the problem in the present analysis and consideration given to the "bulk" properties of the electron avalanches, such as electron mean energy, drift velocity, and the distribution function for the electron energies.

In the absence of a magnetic field, this approach leads (Emeléus, Lunt, and Meek 1936) to the following expression for the first Townsend coefficient,

$$\alpha/p = KW^{-1} \int_0^{\infty} P(V) \cdot V^{\frac{1}{2}} \cdot f(V) \cdot dV, \dots\dots\dots (1.05)$$

where $K = \sqrt{e/150m}$ and is constant,

W is the electron drift velocity,

$P(V)$ is the ionization efficiency of electrons with energy V , at 1 mm Hg pressure, and

$f(V)$ is the distribution function of the electron energies.

In order to extend this approach to the case when a transverse magnetic field is present, it is necessary to determine the influence of the magnetic field on the quantities occurring on the right-hand side of equation (1.05). One may then obtain a new expression for α/p as a function of H/p and E/p .

II. THE INFLUENCE OF A TRANSVERSE MAGNETIC FIELD ON THE ELECTRON AVALANCHE

In order to analyse the influence of the magnetic field on the properties of the electron avalanche, it is necessary to know the variation of the mean free path l with electron velocity u . The assumption usually made is that l is independent of u , and for some gases (notably air) this is approximately true. For hydrogen and helium, however, a more valid assumption is that l is proportional to u (von Engel 1955). Since the values of the Townsend coefficients in static electric fields are best known for hydrogen, experimental work in these laboratories has been carried out using this gas so that, in what follows, the assumption is made that l/u ($=T$, the mean free time) is constant for a given pressure. The error introduced by using this assumption for other gases will be discussed at a later stage.

When only the electric field E is present, any quantity Q under discussion will be denoted by $Q_{0, E/p}$; when both H and E are present, the quantity will be denoted by $Q_{H/p, E/p}$.

(a) *The Electron Drift Velocities, $W_{0, E/p}$ and $W_{H/p, E/p}$*

For the particular case when the mean free time is constant, the drift velocity $W_{0, E/p}$ is given by

$$W_{0, E/p} = EeT/m. \quad \dots\dots\dots (2.01)$$

This expression was first derived by Pidduck (1913) and has been obtained more recently by Davidson (1954) and Huxley (1957*a*), using different methods of derivation.

An expression for the corresponding drift velocity $W_{H/p, E/p}$ in the presence of a transverse magnetic field has been obtained by Huxley (1957*b*) assuming that $l=f(u)$, and, for the particular case when $l/u=T$, his result reduces to

$$W_{H/p, E/p} = \frac{E}{H} \frac{wT}{1+w^2T^2}, \quad \dots\dots\dots (2.02)$$

where $w = He/m$.

This same result can be obtained by a rather different method, which is given in Appendix I, because this particular approach will be used later when discussing the influence of the magnetic field on the secondary coefficient.

From (2.01) and (2.02),

$$W_{0, E/p}/W_{H/p, E/p} = 1 + w^2T^2. \quad \dots\dots\dots (2.03)$$

Also $wT = HeL/mpu$, so that for a given gas

$$w^2T^2 = C(H/p)^2, \quad \dots\dots\dots (2.04)$$

where C is constant $\{=(eL/mu)^2\}$.

Substituting for w^2T^2 in equation (2.03),

$$W_{0, E/p}/W_{H/p, E/p} = 1 + C(H/p)^2. \quad \dots\dots\dots (2.05)$$

(b) *The Velocity Distribution*

The precise form of the velocity distribution of electrons in crossed electric and magnetic fields is not known, except for the particular case when the collisions between electrons and gas molecules are elastic. Allis and Allen (1937) have shown that the distribution function can then be written as

$$f(u) = A_0 u^2 \exp \left\{ -\frac{3m}{M} \left(2 \int \varepsilon/\varepsilon_i^2 d\varepsilon + [H^2/me^2] \int d\varepsilon \right) \right\}, \quad \dots (2.06)$$

where $\varepsilon = \frac{1}{2}mu^2$ and $\varepsilon_i = Eel$.

When the mean free time is independent of u , the integrals in equation (2.06) can be evaluated to give

$$f(u) = A_0 u^2 \exp \left\{ -\frac{3m^3}{2M} \cdot \frac{(1+w^2T^2)}{E^2e^2T^2} \cdot u^2 \right\}, \quad \dots\dots\dots (2.07)$$

so that the distribution of electron velocities is Maxwellian. The distribution is Maxwellian whether the magnetic field is present or not, the effect of the magnetic field being the same as if T were replaced by $T/\sqrt{(1+w^2T^2)}$. Since

T depends only upon the *actual* gas pressure (at a given gas temperature) and $T=L/pu$, the effect of the magnetic field on the distribution function is equivalent to an increase in pressure by the factor $\sqrt{(1+w^2T^2)}$. Thus in this particular case, the magnetic field has no influence on the *form* of the distribution but only reduces the mean electron energy.

Whether this conclusion is valid when inelastic collisions occur remains to be determined. It is of interest in this connexion, however, to note the results of a recent redetermination of theoretical values of $(\alpha/p)_{0, E/p}$ in hydrogen (Blevin and Haydon 1957) showing that for this gas the distribution function may be assumed to be approximately Maxwellian. For hydrogen, then, the "elastic collision" theory cited above seems to fit the experimental results quite well even when inelastic collisions take place. Consequently, it will be assumed that the magnetic field does not affect the *form* of the distribution function, but has the effect of reducing the mean energy.

(c) *The Mean Electron Energy*

If the mean electron energy is \bar{V} when both the magnetic and electric fields are present, then, for equilibrium, the average energy gain per free path must equal the average energy loss at collision, that is,

$$Ey = \lambda(\bar{V}) \cdot \bar{V}, \quad \dots\dots\dots (2.08)$$

where y is the average distance travelled in the direction of the field E between collisions, and $\lambda(\bar{V})$ is the average fractional energy loss at collision when the mean energy is \bar{V} . For a given gas $\lambda(\bar{V})$ depends only upon the distribution of energies, and consequently, using the assumption of the preceding section regarding the distribution function, will be independent of H .

Now

$$\begin{aligned} y &= T \cdot W_{H/p, E/p} \\ &= \frac{L}{pu} W_{H/p, E/p} \quad \dots\dots\dots (2.09) \end{aligned}$$

From (2.08) and (2.09)

$$\frac{E}{p} \cdot \frac{L}{u} \cdot W_{H/p, E/p} = \lambda(\bar{V}) \cdot \bar{V}. \quad \dots\dots\dots (2.10)$$

Since $\lambda(\bar{V})$ is independent of H , $\lambda(\bar{V}) \cdot \bar{V}$ may also be thought of as the average energy loss at collision in a different situation in which there is no magnetic field, and some different value of E/p prevails. Let this value be E/p' .

In this case, the equilibrium condition becomes

$$Ez = \lambda(\bar{V}) \cdot \bar{V}, \quad \dots\dots\dots (2.11)$$

where $z = T \cdot W_{0, E/p'}$ or

$$z = (L/p'u) \cdot W_{0, E/p'} \quad \dots\dots\dots (2.12)$$

From (2.11) and (2.12)

$$(EL/p'u) \cdot W_{0, E/p'} = \lambda(\bar{V}) \cdot \bar{V}. \quad \dots\dots\dots (2.13)$$

Comparing (2.13) and (2.10) shows

$$\begin{aligned} p'/p &= W_{0, E/p'} / W_{H/p, E/p} \dots\dots\dots (2.14) \\ &= \frac{W_{0, E/p'} \cdot W_{0, E/p}}{W_{0, E/p} \cdot W_{H/p, E/p}}. \end{aligned}$$

Now from equation (2.01) $W_{0, E/p'} / W_{0, E/p} = p/p'$, so that

$$p'/p = \sqrt{(W_{0, E/p} / W_{H/p, E/p})}. \dots\dots (2.15)$$

Using (2.05)

$$p' = p \sqrt{\{1 + C(H/p)^2\}}. \dots\dots\dots (2.16)$$

From the definition of p' it follows, then, that in crossed electric and magnetic fields the electrons behave energetically as they would if only the electric field were present, and the pressure were increased from p to p' .

Townsend and Gill (1938) have also derived an expression for the "equivalent pressure" by a consideration of the number of collisions made by an electron when advancing unit distance in the direction of the electric field. Since their result differs from equation (2.16), it is necessary to examine their work more closely.

When there is no magnetic field, the number (n_1) of collisions made by an electron moving unit distance in the direction of E is given by

$$n_1 = (W_{0, E/p} \cdot T)^{-1}. \dots\dots\dots (2.17)$$

With the magnetic field present, the number of collisions becomes

$$n_2 = (W_{H/p, E/p} \cdot T)^{-1},$$

or, using (2.05)

$$n_2 = \{1 + C(H/p)^2\} / (W_{0, E/p} \cdot T). \dots\dots\dots (2.18)$$

Townsend and Gill obtained this result and by comparing (2.17) and (2.18) concluded that the magnetic field has the same influence in this respect as would a decrease in T by the factor $1 + C(H/p)^2$, or, since $T = L/pu$, an increase in pressure by this factor. However, this makes no allowance for the change in the drift velocity which would take place if the pressure were increased.

Using equation (2.01), the result given in (2.17) becomes

$$n_1 = (u/L)p^2/E(e/m)(L/u). \dots\dots\dots (2.19)$$

Similarly (2.18) becomes

$$n_2 = (u/L)p^2\{1 + C(H/p)^2\}/E(e/m)(L/u). \dots\dots (2.20)$$

Comparing equations (2.19) and (2.20) now shows that the magnetic field is identical, in this respect, with an increase in pressure by a factor $\sqrt{\{1 + C(H/p)^2\}}$, in agreement with equation (2.16).

III. CALCULATION OF $(\alpha/p)_{H/p, E/p}$

It is now possible to calculate $(\alpha/p)_{H/p, E/p}$ by way of equation (1.05), that is,

$$(\alpha/p)_{H/p, E/p} = K(W_{H/p, E/p})^{-1} \left\{ \int_0^\infty P(V) \cdot V^{\frac{1}{2}} \cdot f(V) \cdot dV \right\}_{\bar{V}}, \quad \dots (3.01)$$

where the subscript \bar{V} means that the integral is to be evaluated for this mean energy.

It also follows from equation (1.05) and the discussion of Section II (c) that

$$(\alpha/p)_{0, E/p'} = K(W_{0, E/p'})^{-1} \left\{ \int_0^\infty P(V) \cdot V^{\frac{1}{2}} \cdot f(V) \cdot dV \right\}_{\bar{V}}, \quad \dots (3.02)$$

where p' is determined by equation (2.11).

From (3.01) and (3.02)

$$(\alpha/p)_{H/p, E/p} = \frac{W_{0, E/p'}}{W_{H/p, E/p}} \cdot (\alpha/p)_{0, E/p'}, \quad \dots (3.03)$$

or, using (2.14) and (2.16),

$$(\alpha/p)_{H/p, E/p} = \sqrt{\{1 + C(H/p)^2\}} \cdot (\alpha/p)_{0, E/p'}. \quad \dots (3.04)$$

For many gases and over a considerable range of the parameter E/p (von Engel 1955),

$$(\alpha/p)_{0, E/p} = A \exp(-Bp/E), \quad \dots (3.05)$$

so that

$$(\alpha/p)_{0, E/p'} = A \exp[-B(p/E)\sqrt{\{1 + C(H/p)^2\}}]. \quad \dots (3.06)$$

Equations (3.04) and (3.06) then yield

$$(\alpha/p)_{H/p, E/p} = A \sqrt{\{1 + C(H/p)^2\}} \exp[-B(p/E)\sqrt{\{1 + C(H/p)^2\}}]. \quad \dots (3.07)$$

Thus by comparing equations (3.05) and (3.07) it can be seen that the effect on the first Townsend coefficient of the addition of a transverse magnetic field is the same as an increase in pressure by the factor $\sqrt{\{1 + C(H/p)^2\}}$.

Should the empirical relationship given in equation (3.05) be invalid for a particular gas, then equation (3.04) can be used, $(\alpha/p)_{0, E/p'}$ being evaluated from experimental results.

Also equations (3.03) and (2.14) show that

$$(\alpha/p)_{H/p, E/p} = (p'/p)(\alpha/p)_{0, E/p'}.$$

As these two equations depend only upon the assumption that the magnetic field does not alter the form of the distribution function, it would be expected that, if this assumption is valid, the "equivalent pressure" concept is justified for the general case when $L=f(u)$. Following the discussion of Section II (b), however, the assumption that the form of the distribution function is unchanged

depends upon the condition $L/u = \text{constant}$, and it seems unlikely that these two assumptions can be divorced.

It should be emphasized that, in applying the results of the above analysis to pre-breakdown currents in uniform fields, a distinction must be made between the spatial and temporal variations, owing to the fact that the drift velocity, collision frequency, and consequently the electron transit time, are different in the following two systems:

- (i) Magnetic field $= H$,
 electric field $= E$,
 pressure $= p$;
- and (ii) magnetic field $= 0$,
 electric field $= E$,
 pressure $= p'$.

This means that although, as far as the primary ionization is concerned, the same electron multiplication takes place in a given distance, the time taken for this multiplication to be achieved is different in the two systems. This situation originates from the fact that when L/u is constant the average time spent between collisions depends only upon the actual gas pressure.

IV. COMPARISON OF THEORY AND EXPERIMENT

The only measurements of $(\alpha/p)_{H/p, E/p}$ available for comparison with the theory are those made in hydrogen in these laboratories (Blevin 1956) for $50 < E/p < 150 \text{ V cm}^{-1} (\text{mm Hg})^{-1}$, and for $H < 700$ oersteds. The hydrogen used in these experiments was admitted to the vacuum system by diffusion through a palladium thimble. The ionization chamber was not baked out, however, and, as a diffusion pump was not used, the lowest pressure obtainable in the system was about 10^{-3} mm Hg , so that the gas possibly contained small amounts of impurities.

The values of A, B used in (3.07) for the calculation of $(\alpha/p)_{H/p, E/p}$ were 5.6 and 141 respectively, corresponding to the measurements of $(\alpha/p)_{0, E/p}$ made in the same apparatus. These measurements are in good agreement with other recent determinations of $(\alpha/p)_{0, E/p}$ (Blevin, Haydon, and Somerville 1957).

The value of C to be used in equation (3.07) can be determined by several methods.

(i) Since $C = (eL/mu)^2$ and e, m are known, it is only necessary to find L/u . This may be calculated from collision cross-section data such as those determined by Ramsauer (1921) and Brode (1925).

(ii) Microwave measurements of the properties of electric discharges in hydrogen enable the collision frequency $\nu_c = u/L$ (at 1 mm Hg pressure) to be determined (Rose and Brown 1955; Udelson, Creedon, and French 1957), and hence C .

(iii) Equation (2.01) shows that $W_{0, E/p} = \sqrt{C \cdot E/p}$, so that by taking a linear approximation to experimental values of the drift velocity, C and ν_c may be calculated. A summary of determinations of the electron collision frequency ν_c by these methods is given in Table 1.

That the values of v_c obtained by these various methods are different is to be expected, because a collision between an electron and a molecule is defined differently for the different methods (Healey and Reed 1941). For the deter-

TABLE 1
COMPARISON OF COLLISION FREQUENCY MEASUREMENTS

Method	Author	$v_c \times 10^{-9} \text{ sec}^{-1}$
Collision cross-section data ..	Ramsauer	5.9
Collision cross-section data ..	Brode	3.6
Microwave measurements ..	Rose and Brown	4.85
Microwave measurements ..	Udelson, Creedon, and French	4.6
Drift velocities	(As in text)	3.6

TABLE 2
COMPARISON OF EXPERIMENTAL AND THEORETICAL VALUES OF $\alpha_{H/p, E/p}/\alpha_0, E/p$

$E/p=50, p=5.17 \text{ mm Hg}$			$E/p=70, p=5.17 \text{ mm Hg}$		
H/E	$\alpha_{H/p, E/p}/\alpha_0, E/p$		H/E	$\alpha_{H/p, E/p}/\alpha_0, E/p$	
	Expt.	Calc.		Expt.	Calc.
0.5	0.976	0.986	0.5	0.988	0.985
0.8	0.954	0.966	1.0	0.940	0.942
1.0	0.940	0.947	1.5	0.855	0.873
1.5	0.880	0.886	2.0		0.789
2.0	0.815	0.809	3.0		0.597
2.5		0.720			
3.0		0.629			

$E/p=100, p=2.52 \text{ mm Hg}$			$E/p=150, p=2.11 \text{ mm Hg}$		
H/E	$\alpha_{H/p, E/p}/\alpha_0, E/p$		H/E	$\alpha_{H/p, E/p}/\alpha_0, E/p$	
	Expt.	Calc.		Expt.	Calc.
0.5	0.989	0.987	0.5	1.02	1.001
1.0	0.945	0.947	1.0	0.985	0.989
1.5	0.870	0.881	1.5	0.895	0.937
2.0	0.766	0.797	2.0	0.754	0.855
2.5	0.633	0.697	3.0		0.637
3.0		0.591			
4.0		0.405			
5.0		0.260			

mination of the effective cross section a collision is defined as an event in which an electron suffers an appreciable change either in direction of motion or in velocity, whereas for the drift velocity analysis, a collision is defined as an event

in which, on the average, an electron loses all its momentum in any specified direction. Since the constant C has been introduced into the theory by a consideration of the electron drift velocities, it seems appropriate to select $v_c = 3.6 \times 10^9$, or $C = 2.4 \times 10^{-5}$.

Experimental and calculated values of $\alpha_{H/p, E/p} / \alpha_{0, E/p}$ are given in Table 2 for various values of E/p in the range $50 < E/p < 150$. Figure 1 shows the nature of the agreement between the present theory and the observed values, together with the theoretical values of Somerville and Haefner (cf. equation (1.04)), for $E/p = 50$. It can be seen that even at large values of H/E , where the latter

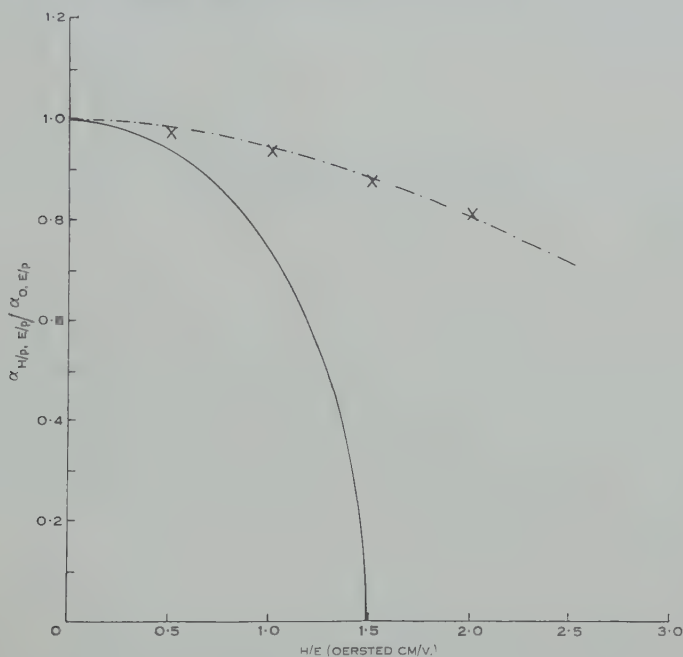


Fig. 1.—Theoretical and experimental values of $\alpha_{H/p, E/p}$ in hydrogen for $E/p = 50 \text{ V cm}^{-1} (\text{mm Hg})^{-1}$. \times Experimental values. — — — Present theory. — Equation (1.04) (cf. Somerville 1952; Haefner 1953).

theory predicts zero values of $(\alpha/p)_{H/p, E/p}$, the present theory is in good agreement with observation. The calculated values are quite sensitive to the value chosen for v_c because v_c^2 appears in the calculation, so that the accurate measurement of $(\alpha/p)_{H/p, E/p}$ in pure hydrogen might well be used as the basis of the determination of v_c in this gas.

V. TOWNSEND'S SECONDARY COEFFICIENT

The secondary processes acting in an electric discharge can be denoted by a generalized coefficient ω/α (Llewellyn-Jones and Parker 1950). When only positive ion and photon action at the cathode are important,

$$\omega/\alpha = \gamma + \delta/\alpha, \quad \dots\dots\dots (5.01)$$

where γ is the average number of electrons liberated from the cathode per incident positive ion and δ is the average number of electrons liberated from the cathode by the photons created in the gas when an electron moves 1 cm in the direction of the electric field. Equation (5.01) is valid, for instance, in low pressure discharges in hydrogen (Morgan 1956), for which gas it has already been shown in Section IV that the present theory for the first Townsend ionization coefficient is applicable.

Equation (5.01) does not take into account, however, the probability that electrons set free by either of these processes may be scattered back to the cathode after a collision with a gas molecule, and be captured there. If k is the fraction of secondary electrons which remain free in the gas, then

$$\omega/\alpha = k(\gamma + \delta/\alpha). \quad \dots\dots\dots (5.02)$$

(a) *The Recapture Coefficient, k*

In the presence of both an electric and a transverse magnetic field there are *two* mechanisms by which electrons leaving the cathode may be returned there and recaptured.

(i) Electrons colliding elastically in the vicinity of the cathode can have sufficient energy to travel against the field E back to the cathode and be recaptured. Let the fraction of electrons which escape recapture by this mechanism be k_1 . If only E is present, then, at those values of E/p where ionization is appreciable, k_1 remains nearly constant (Theobald 1953) with increasing E/p , so that little error is introduced by assuming that k_1 is independent of E/p , and, consequently, of the electron energy. Thus, although the presence of the magnetic field changes the mean electron energy, k_1 can be assumed independent of H for sufficiently high values of E/p .

(ii) There is a second loss mechanism which is not present in the absence of H , namely, loss of secondary electrons which do *not* suffer collision in the gas but return to the cathode under the action of the magnetic field. Let the fraction of electrons which escape recapture by this means be denoted by k_2 . This process has been investigated by Somerville (1952), but slight modification of his theory is required when the mean free time rather than the mean free path is considered to be constant.

If t' is the time required for a secondary electron liberated from the cathode to return there, the number of electrons $N(t')$ travelling for this time without collision is (cf. Appendix I, equation (A2))

$$N(t') = N_0 \exp(-t'/T), \quad \dots\dots\dots (5.03)$$

where N_0 is the number of electrons leaving the cathode. Consequently the fraction of electrons remaining free in the gas is given by

$$1 - N(t')/N_0 = 1 - \exp(-t'/T),$$

and, if r is the probability that an electron will be reflected from the cathode, then from (5.03), $N_0 r \exp(-t'/T)$ electrons will leave the cathode again, and a fraction $r \{\exp(-t'/T)\} \{1 - \exp(-t'/T)\}$ of these will collide and remain free in

the gas. The value of k_2 is obtained by summing these fractions over an infinite number of reflections, giving

$$k_2 = \{1 - \exp(-t'/T)\} / \{1 - r \exp(-t'/T)\}. \quad \dots\dots\dots (5.04)$$

Now if it is assumed that the electrons leave the cathode with zero velocity, the distance x_t travelled in the direction of the electric field is (cf. Appendix I, equation (A1)),

$$x_t = \frac{1}{w} \frac{E}{H} (1 - \cos wt),$$

but, by the definition of t' , $x_t = 0$ when $t = t'$, or

$$t' = 2\pi/w. \quad \dots\dots\dots (5.05)$$

From (5.04) and (5.05)

$$k_2 = \{1 - \exp(-2\pi/wT)\} / \{1 - r \exp(-2\pi/wT)\}. \quad \dots\dots (5.06)$$

The coefficient k in equation (5.02) is then given by

$$k = k_1 k_2. \quad \dots\dots\dots (5.07)$$

(b) *The Variation of γ with Magnetic Field Strength*

For a given gas and cathode surface, γ will depend only upon the energy of the positive ions reaching the cathode. However, the magnetic field has little effect on the motion of positive ions in comparison to the effect on electrons, because of the much greater mass of the ions. Equation (2.16) shows (when C is evaluated for positive ions) that the "equivalent pressure" for the ions is very little different from the actual pressure, except for large values of H/p . With this restriction,

$$\gamma_{H/p, E/p} = \gamma_{0, E/p}. \quad \dots\dots\dots (5.08)$$

The limiting value of H/p for which this is valid must be evaluated for each gas.

(c) *The Variation of δ/α with Magnetic Field Strength*

At low pressures when photon absorption in the gas is negligibly small, δ can be written in the form

$$\delta = \eta \theta g, \quad \dots\dots\dots (5.09)$$

where θ is the average number of photons produced by an electron moving 1 cm in the direction of the electric field,

g is a geometrical factor determining the probability that a photon will reach the cathode, and

η is the probability that a photon reaching the cathode will liberate an electron.

By analogy with equation (1.05), the excitation coefficient θ is given by

$$(\theta/p)_{0, E/p} = K(W_{0, E/p})^{-1} \int_0^\infty P'(V) \cdot V^{\frac{1}{2}} \cdot f(V) \cdot dV, \quad \dots (5.10)$$

where $P'(V)$ is the efficiency of excitation. Proceeding in the same manner as for ionizing processes, it follows that

$$(\theta/p)_{H/p, E/p} = (p'/p)(\theta/p)_{0, E/p}. \quad \dots\dots\dots (5.11)$$

Now η will depend upon the energy of the incident photon and, when there are photons of different energies present, upon the relative abundance of photons in each energy group.

When excitations occur to different energy levels, $P'(V)$ in equation (5.10) can be replaced by $P'_1(V) + P'_2(V) + \dots + P'_n(V)$, where $P'_n(V)$ is the excitation efficiency for the n th level. Thus, if $(\theta_j)_{0, E/p}$ is the number of photons produced which have energies characterized by the j th level of the gas molecule, then

$$(\theta_j/p)_{0, E/p} = K(W_{0, E/p})^{-1} \int_0^\infty P'_j(V) \cdot V^{\frac{1}{2}} \cdot f(V) \cdot dV,$$

and similarly,

$$(\theta_k/p)_{0, E/p} = K(W_{0, E/p})^{-1} \int_0^\infty P'_k(V) \cdot V^{\frac{1}{2}} \cdot f(V) \cdot dV,$$

so that

$$(\theta_j/\theta_k)_{0, E/p} = \int_0^\infty P'_j(V) \cdot V^{\frac{1}{2}} \cdot f(V) \cdot dV / \int_0^\infty P'_k(V) \cdot V^{\frac{1}{2}} \cdot f(V) \cdot dV. \quad \dots\dots (5.12)$$

By a similar procedure, the ratio $(\theta_j/\theta_k)_{H/p, E/p}$ can be found, giving an equation similar to (5.12) but in which the integrals on the right-hand side must be evaluated for a different mean energy. It has already been shown that this mean energy corresponds to the case when there is only the electric field present, but the pressure is increased from p to p' . It follows, therefore, that,

$$(\theta_j/\theta_k)_{H/p, E/p} = (\theta_j/\theta_k)_{0, E/p'}. \quad \dots\dots\dots (5.13)$$

This result means physically that the relative abundance of photons in the energy groups is the same in the presence of both E and H , as in the case when only E is present, provided that the pressure is increased from p to p' in the latter case, that is,

$$\eta_{H/p, E/p} = \eta_{0, E/p'}. \quad \dots\dots\dots (5.14)$$

The value of g for a given electrode configuration depends only upon the manner in which the production of photons is distributed throughout the discharge space. In crossed fields, the number $\Delta n(x_1)$ of photons produced by N electrons moving a distance Δx in the direction of the electric field, having already travelled a distance x_1 from the cathode, is

$$\Delta n(x_1) = N \cdot \theta_{H/p, E/p} \cdot \Delta x.$$

If N_0 electrons originally leave the cathode, then

$$N = N_0 \cdot \exp \left\{ \int_0^{x_1} \alpha_{H/p, E/p} \cdot dx \right\},$$

and

$$\Delta n(x_1) = N_0 \cdot \theta_{H/p, E/p} \cdot \exp \left\{ \int_0^{x_1} \alpha_{H/p, E/p} \cdot dx \right\} \cdot \Delta x.$$

Similarly the number of photons produced by electrons moving a distance Δx at a distance x_2 from the cathode is

$$\Delta n(x_2) = N_0 \cdot \theta_{H/p, E/p} \cdot \exp \left\{ \int_0^{x_2} \alpha_{H/p, E/p} \cdot dx \right\} \cdot \Delta x,$$

so that

$$\Delta n(x_1)/\Delta n(x_2) = \exp \left\{ \int_{x_2}^{x_1} \alpha_{H/p, E/p} dx \right\}. \quad \dots\dots\dots (5.15)$$

From (2.16), (3.04), and (5.15)

$$\Delta n(x_1)/\Delta n(x_2) = \exp \left\{ \int_{x_2}^{x_1} \alpha_{0, E/p} dx \right\},$$

and it follows that the distribution of photon production in the gap when both E and H are present is the same as when only E is acting, but the pressure is increased from p to p' , that is,

$$g_{H/p, E/p} = g_{0, E/p'}. \quad \dots\dots\dots (5.16)$$

Combining equations (2.16), (3.04), (5.09), (5.11), (5.14), and (5.16),

$$(\delta/\alpha)_{H/p, E/p} = (\delta/\alpha)_{0, E/p'}, \quad \dots\dots\dots (5.17)$$

or, using (5.02), (5.07), (5.08), and (5.17),

$$(\omega/\alpha)_{0, E/p} = k_1 \{ \gamma_{0, E/p} + (\delta/\alpha)_{0, E/p'} \}, \quad \dots\dots\dots (5.18)$$

and

$$(\omega/\alpha)_{H/p, E/p} = k_1 k_2 \{ \gamma_{0, E/p} + (\delta/\alpha)_{0, E/p'} \}. \quad \dots\dots (5.19)$$

VI. THE SECONDARY COEFFICIENT FOR SMALL H/p

Equation (5.06) shows that for small values of H/p the recapture coefficient k_2 is very close to unity, so that equations (5.18) and (5.19) can be written as

$$(\omega/\alpha)_{0, E/p} - (\omega/\alpha)_{H/p, E/p} = k_1 \{ (\delta/\alpha)_{0, E/p} - (\delta/\alpha)_{0, E/p'} \}. \quad \dots\dots (6.01)$$

Furthermore, if the fraction of the secondary coefficient due to photons is known at any value of E/p , that is, the ratio $k_1(\delta/\omega)_{0, E/p} = f$ (say) is known, then equation (6.01) may be rewritten

$$(\omega/\alpha)_{H/p, E/p} - (1-f)(\omega/\alpha)_{0, E/p} = k_1(\delta/\alpha)_{0, E/p'}.$$

It follows then that if a value is assumed for f at a given E/p (obtained, for instance, from measurements of the formative time lag of breakdown (Morgan 1956)) it is possible, by measuring the total secondary coefficient ω/α with and without a transverse magnetic field present, to determine the photon contribution $k_1(\delta/\alpha)_{0, E/p'}$ at a value E/p' . By repeating this procedure the actual contribution

due to photons can be determined at any lower value of E/p . In this way the fractional contributions obtained by measurements of the formative time lag of breakdown, in the absence of a magnetic field, may be checked.

VII. CONCLUSIONS

The above theoretical investigation has shown that the presence of a transverse magnetic field has the same effect on many of the "bulk properties" of an electron swarm in hydrogen as would an increase in the gas pressure. This equivalent increase in pressure has been determined and used to evaluate a theoretical expression for the first Townsend ionization coefficient in hydrogen which has been shown to be in good agreement with recent experimental determinations.

TABLE 3
VALIDITY OF EQUIVALENT PRESSURE CONCEPT

Property Associated with Electron Avalanche	Actual Situation		Equivalent Situation	
	Transverse magnetic field = H		Transverse magnetic field = 0	
	Uniform electric field = E		Uniform electric field = E	
	Gas pressure = p		Gas pressure = p'	
Distribution function	$f(V)$		$f(V)$	
First Townsend coefficient	α		α	
Collisions/cm of drift	n		n	
Excitation coefficient	θ		θ	
Photo-emission probability	η		η	
Geometrical factor	g		g	
Drift velocity	W		$(p'/p)W$	
Collision frequency	ν_c		$(p'/p)\nu_c$	

The detailed discussion of the secondary coefficient ω/α has also shown that, for small magnetic fields, the equivalent pressure concept is applicable to the influence of H on the photon contribution. It is not applicable to the influence of small magnetic fields on the liberation of electrons by positive ion bombardment of the cathode. As the magnetic field increases, the recapture of secondary electrons by the cathode becomes increasingly important, so that ω/α is diminished owing to this effect.

The extent to which the equivalent pressure concept developed in this analysis is valid is summarized in Table 3. It may be seen that the concept may be validly applied to all aspects of the spatial growth of currents which depend only on the first six quantities listed in the table. However, owing to the fact that the electron transit time is not identical in the two systems, an analysis of the *temporal* growth of pre-breakdown currents cannot be treated in terms of the particular value of the equivalent pressure derived in this analysis.

A further paper relating the present theory to breakdown characteristics in crossed fields is being prepared.

VIII. ACKNOWLEDGMENTS

This problem was suggested by Professor J. M. Somerville. The authors wish to acknowledge this suggestion and to thank him for his helpful advice and discussion throughout the course of this investigation.

IX. REFERENCES

- ALLIS, W. P., and ALLEN, H. W. (1937).—*Phys. Rev.* **52**: 703.
 BLEVIN, H. A. (1956).—M.Sc. Thesis, University of New England.
 BLEVIN, H. A., and HAYDON, S. C. (1957).—*Aust. J. Phys.* **10**: 590.
 BLEVIN, H. A., HAYDON, S. C., and SOMERVILLE, J. M. (1957).—*Nature* **179**: 38.
 BRODE, R. B. (1925).—*Phys. Rev.* **25**: 636.
 DAVIDSON, P. M. (1954).—*Proc. Phys. Soc. Lond.* **B 67**: 159.
 EMELÉUS, K. G., LUNT, R. W., and MEEK, C. M. (1936).—*Proc. Roy. Soc. A* **156**: 394.
 VON ENGEL, A. (1955).—"Ionized Gases." (Clarendon Press: Oxford.)
 HAEFFER, R. (1953).—*Acta Phys. Austr.* **7**: 52.
 HEALEY, R. H., and REED, J. W. (1941).—"The Behaviour of Slow Electrons in Gases." (Amalgamated Wireless (Australasia) Ltd.: Sydney.)
 HUXLEY, L. G. H. (1957a).—*Aust. J. Phys.* **10**: 118.
 HUXLEY, L. G. H. (1957b).—*Aust. J. Phys.* **10**: 240.
 LLEWELLYN-JONES, F., and PARKER, A. B. (1950).—*Nature* **165**: 960.
 MORGAN, C. G. (1956).—*Phys. Rev.* **104**: 566.
 PIDDUCK, F. B. (1913).—*Proc. Roy. Soc. A* **88**: 296.
 RAMSAUER, C. (1921).—*Ann. Phys., Lpz.* **64**: 513.
 ROSE, D. J., and BROWN, S. C. (1955).—*Phys. Rev.* **98**: 310.
 SOMERVILLE, J. M. (1952).—*Proc. Phys. Soc. Lond.* **B 65**: 620.
 THEOBALD, J. K. (1953).—*J. Appl. Phys.* **24**: 123.
 TOWNSEND, J. S., and GILL, E. W. B. (1938).—*Phil. Mag.* **26**: 290.
 UDELSON, B. J., CREEDON, J. E., and FRENCH, J. C. (1957).—*J. Appl. Phys.* **28**: 717.
 VALLE, G. (1950).—*Nuovo Cimento* **7**: 174.
 WEHRLI, M. (1922).—*Ann. Phys., Lpz.* **69**: 285.

APPENDIX I

The Drift Velocity of Electrons in a Transverse Magnetic Field

Consider an electron moving in a gas, at pressure p mm Hg, under the influence of a uniform electric field E in the $0x$ direction and a uniform magnetic field H in the $0z$ direction.

Putting $He/m=w$, the equation of motion for the electron is (see e.g. Healey and Reed 1941)

$$x_t = (1/w)[\{E/H - u_y\}\{1 - \cos wt\} + u_x \sin wt], \quad \dots \quad (\text{A1})$$

where x_t is the distance travelled in the $0x$ direction in time t , and u_x, u_y are the initial velocities in the $0x, 0y$ directions.

Now, if the mean free time T is independent of u , then the number (dN) of collisions made by N electrons in a time interval dt , is

$$dN = -N(dt/T),$$

or

$$N = N_0 \exp(-t/T), \quad \dots \quad (\text{A2})$$

where N_0 is the total number of electrons.

If a collision is defined as an event in which, on the average, an electron loses all of its momentum in any specified direction, then averaging for initial velocities in (A1) gives

$$\bar{x}_t = (E/Hw)(1 - \cos wt).$$

Averaging over free times, and using (A2),

$$\bar{x} = (E/H)wT^2/(1 + w^2T^2),$$

or

$$W_{H/p, E/p} = \bar{x}/T = (E/H)wT/(1 + w^2T^2).$$

HIGH RESOLUTION CINEMATOGRAPHY OF THE SOLAR PHOTOSPHERE

By R. E. LOUGHHEAD* and V. R. BURGESS*

[*Manuscript received September 10, 1957*]

Summary

A description is given of a new photoheliograph designed for time-lapse cinematography on 35 mm film of any selected portion of a 20 cm solar image. The interval between successive photographs can be varied from 5 to 120 sec. Air-suction devices have been incorporated to suppress bad seeing originating within the telescope. The performance of the telescope is briefly discussed and the advantages of the cinematographic technique indicated. A discussion of the effects of atmospheric seeing is given in order to emphasize the great caution needed in the interpretation of solar photographs.

I. INTRODUCTION

The improvements over the past 50 years in spectroscopic and spectroheliographic methods of observing the Sun have enormously increased our knowledge of the Sun's outer layers, but have not been paralleled by corresponding advances in the direct photography of the solar photosphere. Most of the telescopes used in spectroscopic studies during this period have been "tower-type" instruments, which, however, are by their very nature somewhat handicapped in delineating the finest details on the Sun's surface. Even under conditions of excellent atmospheric seeing the actual performance of such instruments seldom attains the theoretical limit owing to disturbing air currents throughout the optical paths, thermal distortion of the mirror optics, and unequal heating inside and outside the dome or tower. To date most good photographs of the photosphere have in fact been taken with refractors of various types. The quality of the plates obtained by Janssen (1877) probably remains unsurpassed, but excellent results have been obtained by Hansky (1908) and Chevalier (1912) and more recently by a number of other workers. Most of the later investigators have avoided the introduction of enlarging cameras (cf. Keenan 1953) for fear that image distortions may result from heated air in the neighbourhood of the enlarging lens; this, however, has entailed the use of telescopes of considerable focal length to obtain the size of image needed for high resolution photography.

In the present paper we describe a 5 in. photoheliograph which is accommodated conveniently on a mounting of moderate size, but which nevertheless effectively provides a solar image 20 cm in diameter. Essentially the instrument is a 5 in. refractor designed for time-lapse cinematography of any selected portion of the solar image on 35 mm film. The interval between successive photographs

* Division of Physics, C.S.I.R.O., University Grounds, Chippendale, N.S.W.

can be varied from 5 to 120 sec. The smearing effect on the image due to atmospheric seeing is minimized by taking exposures of about 1 msec duration. In addition, devices have been incorporated to ensure that the whole telescope is shielded as much as possible from the Sun and that heated air formed at those surfaces unavoidably exposed is removed by an air-suction technique. The purpose of these devices is to reduce bad seeing originating within the telescope to negligible proportions, and thus enable full advantage to be taken of the available optical resolution during times of good atmospheric seeing.

II. DESIGN

(a) Mounting

The photoheliograph is mounted in the open air on one face of a 10 ft "equatorial spar" at the C.S.I.R.O. Division of Physics Solar Observatory, located in flat country some 30 miles south-west of Sydney at an altitude of about 200 ft. The equatorial spar consists of a rectangular box with sides 16.5 by 12.5 in. made of $\frac{1}{4}$ in. steel plate, whose three clear faces provide mountings for instruments. A general view of the instrument and mounting is shown in Plate 1. The spar is made to follow the Sun by means of a synchronous motor driving in hour angle and a photoelectric guider fed by an auxiliary 3.5 in. telescope. This produces a solar image on an occulting disk behind which are placed four photoelectric cells. Servomotors providing motion in hour angle and declination are driven by the amplified signals from the photocells. Guiding to 1-2 sec of arc can be achieved under the best conditions. The guider and spar are based on designs kindly supplied by Dr. W. O. Roberts of the High Altitude Observatory, Climax, Colorado.

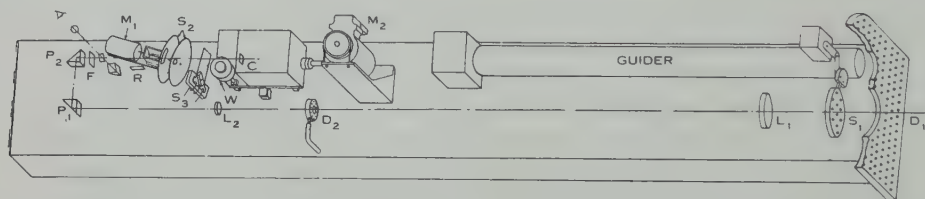


Fig. 1.—Layout of the photoheliograph. *C*, camera; *D*₁, front shield; *D*₂, diaphragm at primary focus; *L*₁, objective; *L*₂, magnifying lens; *M*₁, shutter motor; *M*₂, camera motor; *P*₁, *P*₂, prisms; *R*, reflex mirror; *S*₁, objective shutter; *S*₂, sector-disk shutter; *S*₃, blade shutter; *W*, clock.

(b) Optical System

Figure 1 gives a schematic drawing of the layout of the photoheliograph. *L*₁ is a 5 in. objective which produces a 16 mm image of the Sun on a small diaphragm *D*₂. This lens is a cemented doublet mounted free of strain in a cell designed by Coulman and Norton (unpublished data). The portion of the image corresponding to the aperture in *D*₂ is magnified by a second lens *L*₂ and the image is formed at the gate of a 35 mm Debrie cinecamera *C*. Prisms *P*₁ and *P*₂ turn the light beam around the end of the spar, thus permitting the use of a high magnification in a relatively confined space. Alternative magnifying lenses *L*₂ are provided to give effective image diameters at *C* of 8.5 and 20 cm.

L_2 is provided with graduated screws giving motions in the north-south and east-west directions, so that any desired region of the Sun can be brought onto the camera gate. The optical system is designed so that off-axis aberrations thereby introduced do not exceed the Rayleigh tolerance. R is a reflex mirror used in focusing the telescope and F is a glass filter of bandwidth 800 \AA centred at 5400 \AA . At this wavelength the theoretical limit of resolution is about 0.8 sec of arc .

The film used is Recordak Micro-File, a fine-grained emulsion, which is developed for about 6 min in D19. Laboratory tests on this film at different levels of object contrast have shown that the performance of the telescope at the higher of the two magnifications is not limited by film resolution.

(c) *Suction Devices*

By the use of suction devices, heated air formed at surfaces exposed to the Sun is removed and replaced by air at the ambient temperature. The photo-heliograph and the front of the equatorial spar are shielded by a hollow aluminium diaphragm D_1 . The suction system is used to draw air through numerous holes in the front surface of D_1 , which is thus kept at the ambient temperature. S_1 is a solenoid-operated shutter which excludes light from the telescope except for a short time when an exposure is due. Unwanted light at the primary focus is reflected away by the diaphragm D_2 , which also reduces scattered light in the optical system. Both S_1 and D_2 are kept at ambient temperature by air flow in the same way as D_1 . The suction system is also used to remove any heat generated by the solenoid operating S_1 and by the electric motor M_1 driving the rotating sector-disk shutter S_2 . The reduced pressure in the suction system is maintained by a $\frac{1}{2}$ h.p. electric forge-blower, which is situated some 15 ft from the telescope and connected to the suction devices by an underground pipe and flexible couplings.

The air-suction devices are based on the results of laboratory experiments. Plate 2 shows two interferograms obtained with light passing over a perforated metal surface, heated by radiation to about the same extent as by ordinary sunlight. The distortion suffered by the wave-front in passing over the irradiated surface is shown in Plate 2, Figure 1, while Plate 2, Figure 2, illustrates the suppression of this effect when air is drawn through the perforated area. In practice the quantity of air drawn through the suction devices in the telescope is sufficient to reduce the temperature of the irradiated surfaces to the ambient value within a few minutes, even after exposure to sunlight for several hours.

(d) *Shutter Unit*

The focal plane shutter mounted in front of the camera consists of a solenoid-operated blade shutter S_3 and a rotating sector-disk shutter S_2 . The latter consists of two coaxial sector-disks which are driven at a constant relative speed by an electric motor. The high speed disk controls the exposure time; with a sector angle of 4° the effective exposure is of the order of 1 msec. The second disk rotates at one-sixth the speed of the first and therefore, in the absence of a blade shutter, would permit an exposure every sixth rotation of the fast disk.

The purpose of this arrangement is to give the blade shutter S_3 sufficient time in which to operate, and thus to sequence the exposures made by the telescope.

To record photospheric detail of low contrast, the film must be developed to a high gamma; therefore some form of exposure control is necessary to avoid large changes in film density due to changing atmospheric transparency and solar zenith distance. The exposure control unit is located on the lower face of the spar just beneath the 5 in. objective (cf. Plate 1). It consists of a photocell unit fed with sunlight by a small auxiliary telescope. The signal from the photocell, after amplification, controls the speed of the motor driving the rotating sector-disk shutter S_2 in such a way that the product of the light flux and exposure time remains constant. Laboratory tests indicate a constancy to within 1 per cent. over extended periods. The circuit is designed to provide stability against changes either in mechanical load or power supply voltage. A fuller account of this device will be published by one of us elsewhere.

(e) Programme Controller

The photoheliograph is designed to take photographs automatically on a cycle which can be varied from 5 to 120 sec. The automatic control is obtained by means of a timing device providing electric pulses at regular intervals to drive a slave unisector in the programme control unit. This unisector actuates a number of relays which, in conjunction with a commutator mounted on the slower shaft of the sector-disk shutter S_2 , control the operations of the devices in the photoheliograph involved in taking an exposure. These operations are as follows:

- (1) the camera motor (M_2 in Fig. 1) winds the film on one frame,
- (2) the objective shutter S_1 opens,
- (3) the blade shutter S_3 opens,
- (4) the rotating shutter S_2 makes the exposure,
- (5) S_1 and S_3 close.

The times at which S_3 is opened and closed are controlled by the commutator on the shaft of the slower sector of S_2 ; this eliminates multiple exposures. The image of a clock (W in Fig. 1), together with the date, is recorded in a corner of each photograph. The programme controller ensures that the clock is illuminated within a fraction of a second of the actual exposure on the Sun.

III. PERFORMANCE

Plate 3 is an example showing the quality of the best photographs obtained to date. Although certain regions are slightly affected by seeing, the granulation is well resolved and other detail of the order of 1 sec of arc is clearly visible; this detail approaches the theoretical limit of resolution of the 5 in. objective. Series of photographs of this quality have been obtained even after the telescope has been exposed to the heating effect of the Sun for several hours. This fact illustrates the value of suction devices in reducing seeing in or near the telescope.

The quality of the usable photographs among the 1800 or so on a given film is visually assessed on a scale of 1 to 5. Although such an assessment is naturally

somewhat subjective, it is useful in giving some idea of the proportion of good to bad photographs when the cinematographic technique is employed. Quality 1 refers to photographs comparable with Plate 3; quality 5 refers to photographs which resolve the granulation but are badly affected by seeing. Photographs of quality 3 or worse often show the familiar *réseaux* believed by some early workers to be a real solar feature. However, examination of neighbouring photographs, taken at 5-sec intervals, confirms the conclusion of Rosch (1955) that the phenomenon is due to poor seeing. Photographs of quality 5 are of value in studying small structures only when a number of other photographs of sufficient quality, taken about the same time, are available. On the best film taken to date the percentages of photographs of qualities 1 to 5 are respectively 1, 2, 3, 8, and 11 per cent. Thus even on a good film the number of usable photographs is a relatively small fraction of the total. On the other hand, if still photography were employed instead of cinematography, the observer would be fortunate to obtain any good photographs at all under comparable atmospheric conditions.

The study of good quality films obtained with this equipment has shown that it is essential to allow for the effects of atmospheric seeing when studying changes in photospheric structures (cf. Bray and Loughhead 1957). For example, a small bright point of light visible on a good quality photograph might appear on a later photograph of poorer quality as a faint diffuse patch or even fail to appear at all. The misleading impression due to poor seeing is enhanced by the non-linear response of the emulsion to fluctuating light intensities. Moreover, the pattern of seeing on a poor photograph is so complex that, while a given fine structure might be completely obliterated, neighbouring detail of comparable size may be practically unaffected. Finally, quite large structures, many times bigger than the smallest resolvable detail on a given photograph, can be distorted by seeing to a surprising extent.

One advantage of the cinematographic technique is that many photographs can be taken in a time short compared with the time scale of the phenomenon under study, and thus by examining a number of photographs one can detect spurious effects due to seeing. An awareness of the effects of seeing is a prerequisite to the successful interpretation of high resolution films of the photosphere.

IV. ACKNOWLEDGMENTS

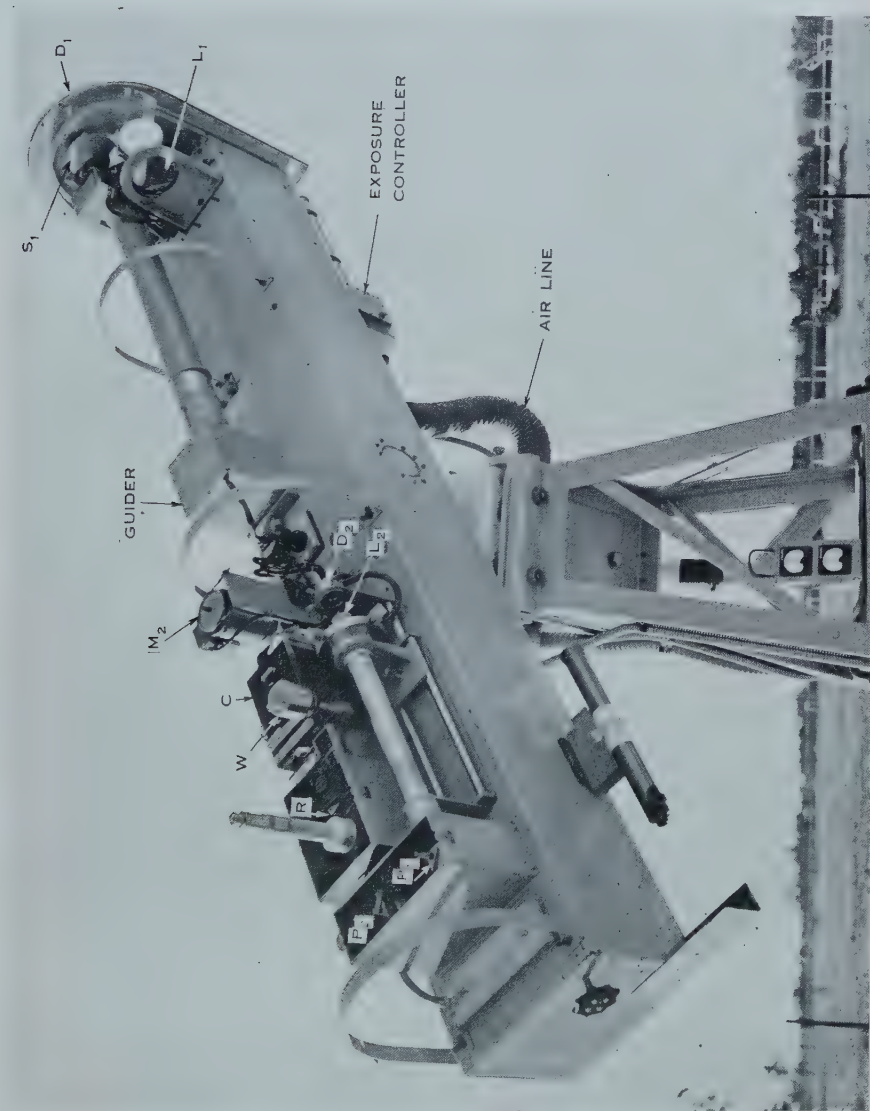
This project was originated by Dr. R. G. Giovanelli and carried out under his general direction. The success of this project is due largely to the enthusiastic cooperation and assistance afforded by many of our colleagues in the Division of Physics, C.S.I.R.O. We are indebted to Mr. C. Coulman for the design and testing of the optical system, and to Dr. R. J. Bray for his assistance in the final stages of the installation and for his critical discussion of this manuscript. Section III of this paper is based on the results of work undertaken jointly by one of us (R.E.L.) and Dr. R. J. Bray. Our thanks are due to Mr. R. Abell for the figuring of the 5 in. objective and the construction of other optical components, to Mr. G. Norton for technical advice and the mounting of the 5 in. objective, to Mr. K.

Regel for undertaking the major part of the mechanical construction, and to Mr. A. Young for supervising the construction of the electronic guider. We also wish to thank Messrs. P. Howden, T. J. Nagel, and R. Rabbige for their work on the electrical installation of the equipment. Members of the Drawing Office and Workshop Staffs at the National Standards Laboratory, Sydney, rendered valuable service in the design and mechanical construction of the equipment.

V. REFERENCES

- BRAY, R. J., and LOUGHHEAD, R. E. (1957).—*Observatory* **77** : 201.
CHEVALIER, P. S. (1912).—*Ann. Obs. Zé-Sé* Vol. 8c.
HANSKY, A. (1908).—*Mitt. Pulkovo Obs.* **3** : 1.
JANSSEN, J. (1877).—*C.R. Acad. Sci., Paris* **85** : 373.
KEENAN, P. C. (1953).—"The Sun." pp. 597-600. (Univ. Chicago Press.)
ROSCH, J. (1955).—*C.R. Acad. Sci., Paris* **240** : 1630.

HIGH RESOLUTION CINEMATOGRAPHY OF THE SOLAR PHOTOSPHERE



Photoheliograph and mounting. *C*, camera; *D*₁, diaphragm at primary focus; *L*₁, objective; *L*₂, magnifying lens; *M*₂, camera motor; *P*₁, *P*₂, prisms; *R*, reflex mirror; *S*₁, objective shutter; *W*, clock.

HIGH RESOLUTION CINEMATOGRAPHY OF THE SOLAR PHOTOSPHERE



Fig. 1.—Interferogram showing the distortion of the wave-front of a light beam passing over a heated metal surface. The heating effect of sunlight on the surface is simulated by artificial radiation.

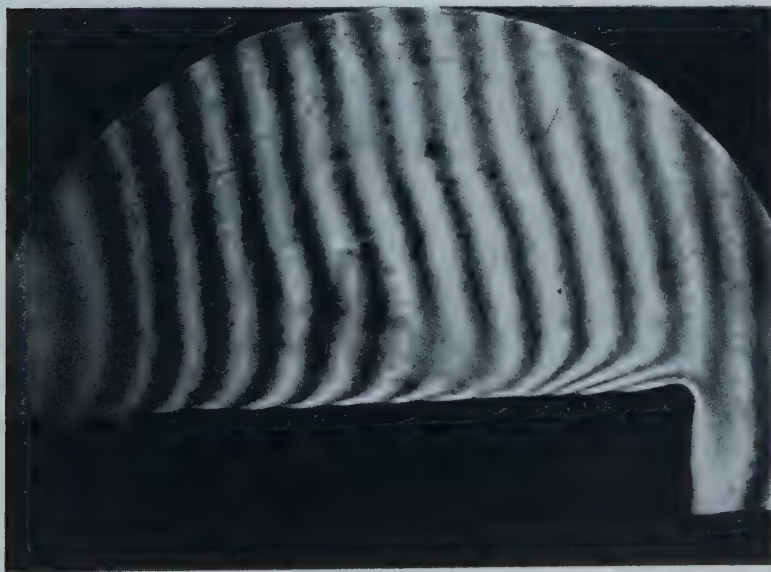
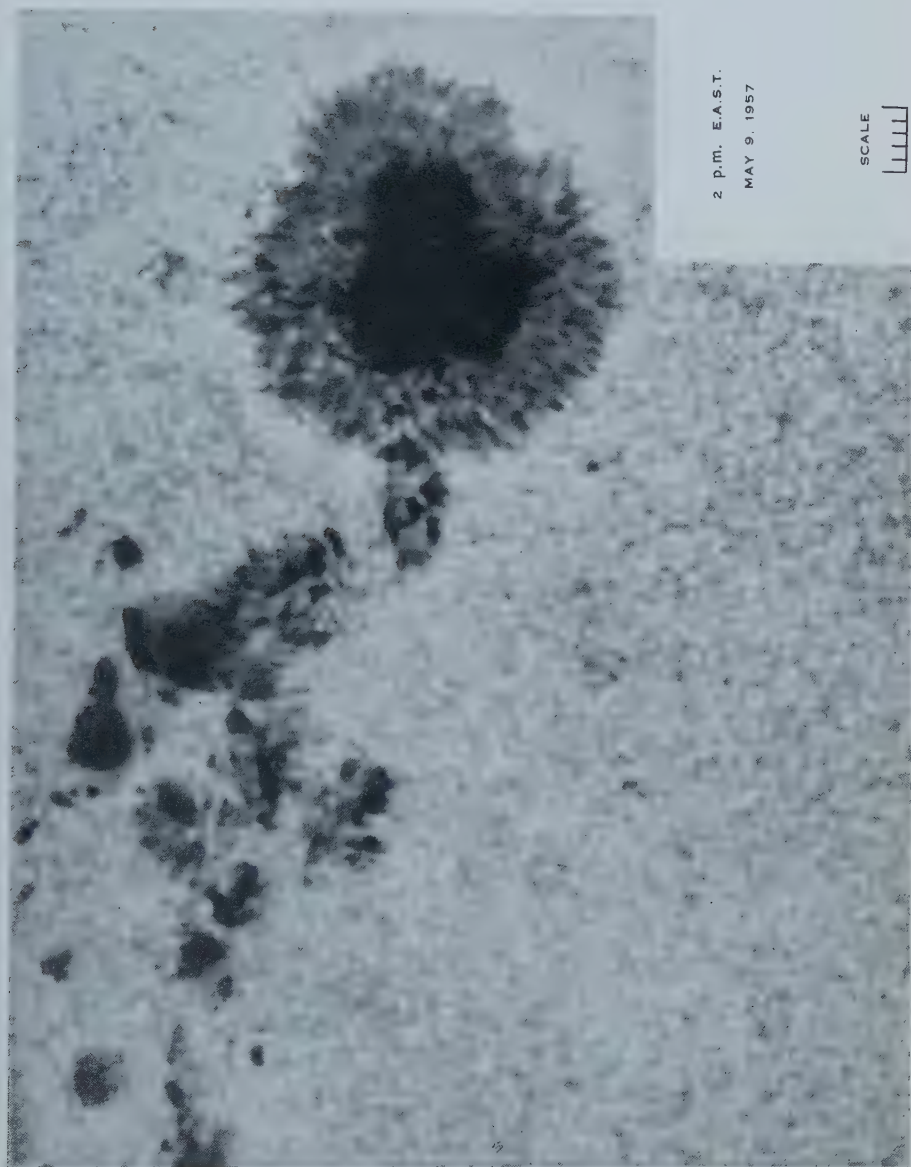


Fig. 2.—Interferogram showing the reduction of the distortion when air is drawn through holes in the heated surface.

HIGH RESOLUTION CINEMATOGRAPHY OF THE SOLAR PHOTOSPHERE



Enlargement of a photograph taken with the photoheliograph on 35 mm film.

THE DESIGN OF PHOTOGRAPHIC OBJECTIVES OF THE TRIPLET FAMILY

I. THE DESIGN OF THE TRIPLET TYPE 111 OBJECTIVE

By F. D. CRUICKSHANK*

[*Manuscript received October 16, 1957*]

Summary

A general study has been made of the initial design of photographic objectives of the triplet family of objectives. A classification is suggested. The point of view is taken that it is convenient to regard the first two members of a triplet as constituting a compound corrector system to the rear positive member.

Part I. An account is given of the design of the type 111 triplet objective to which any triplet may be reduced. A simple algebraic solution is then possible for the arrangement of the initial thin component system having any selected corrector power, χ . It is shown that in triplets of unit power the only systems of any practical interest have corrector powers approximately in the range $-2 < \chi < 0.4$. In this range χ specifies the distribution of power between the two positive members of the triplet and is the variable which determines the spherical aberration when all other primary aberrations have been adjusted to desired small values. The effect of the residual values of the Petzval curvature and the chromatic aberration on the initial solution is discussed. The method provides a starting point for the development of the design of any objective having a basic triplet structure.

I. INTRODUCTION

It is proposed to consider the general problem of the initial design of photographic objectives which are basically of a triplet construction. For this purpose it is convenient to regard these as constituting a family of lens systems logically derived from the simple triplet, the process of derivation consisting of the replacement of one or more members of the parent triplet by a group of lenses. The replacing group is commonly a cemented doublet or triplet. Some of the members of this family are very well known and are currently manufactured under various names such as the Tessar, the Heliar, the Sonnar, etc. In Figure 1 some of the possible constructions within the family are shown, the arrangement being determined by the complexity of the components. In this scheme it is convenient to assign to each construction a type number based on the number of elements in the three component groups. Thus the Tessar is specified as a type (1, 1, 2) triplet or, more briefly, type 112. On the same basis the Heliar is a type 212 triplet, the Sonnar is a type 133 triplet, and so on. The broad problem then is to establish a general pattern according to which the design of any member of this family of objectives may be developed from first principles.

* Department of Physics, University of Tasmania, Hobart.

The process of designing an optical system falls broadly into three stages. In the first of these it is usual to determine an initial arrangement of thin lenses in which certain basic requirements such as power, Petzval sum, and paraxial achromatism are met. In the second stage the requisite axial thicknesses are introduced and the shapes of the lenses determined so that the primary (third order) aberrations of the system are under control. In the third stage the secondary (fifth order) and higher order aberrations are determined and some means of differential correction employed for the final adjustment of the system until the desired balance of aberrations is achieved. An early problem then is

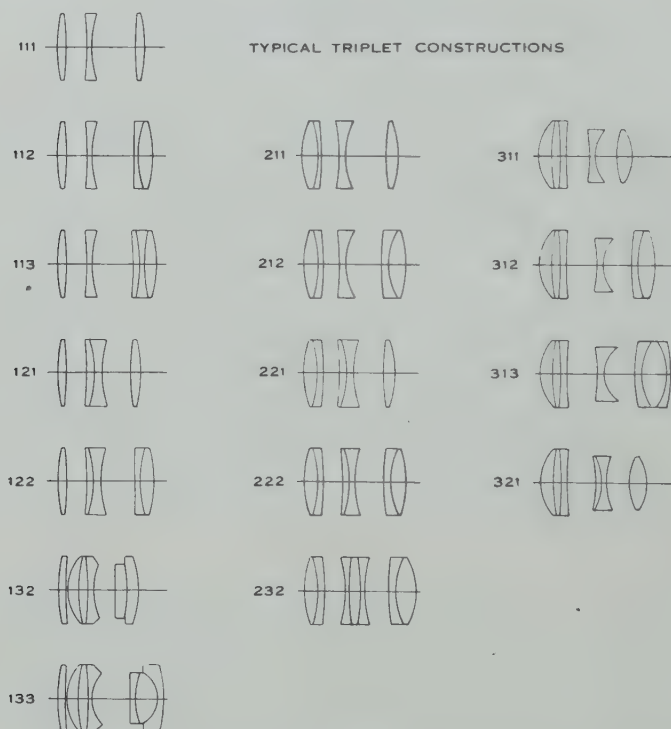


Fig. 1.—Some members of the triplet family of objectives.

to obtain the initial thin lens arrangement in the case of the parent triplet, type 111. This is discussed in detail in the opening sections of the present paper (Part I). This must be in a form capable of application to the more complex members of the family of objectives. Any other member of the family may then be considered as derived from a type 111 triplet, the derivation involving the replacement of one or more of the thin singlet components of the parent type by thin doublets or triplets.

In Part II (in preparation) this process of replacement is discussed. Consider the thin lens (ϕ , N , V) of power ϕ made from glass of refractive index N and V -number V . Suppose this is to be replaced by a doublet consisting of two thin lenses (ϕ_1 , N_1 , V_1), ($k\phi_1$, N_2 , V_2) in contact. It will be obvious that this can be

done in such a way that the doublet has the same power and *either* the same Petzval sum *or* the same paraxial chromatic aberrations as the singlet lens which it replaces. Conversely, such a doublet constructed from some pair of glasses is equivalent as regards power, Petzval sum, and chromatic aberrations to a singlet lens having glass constants (N , V) which need not correspond with those of any known glass. These values (N , V) may be varied widely by intelligent choice of the two glasses of the doublet and the value of k . Again, when the replacement of a single lens by a cemented triplet is considered it will be seen that it is formally possible to design a thin cemented triplet which is equivalent as regards power, Petzval sum, and chromatic aberrations to a single thin lens having any desired value of N and, independently thereof, any desired value of V .

From the point of view of the first stage of design, then, the passage from the basic type 111 triplet to any other member of the family is equivalent simply to a change of glass. In other words, in considering the initial design of the basic type 111 the glass constants of the thin lens components may be treated as continuously variable within a certain range. Values which do not correspond to known glass types may be achieved by using compound components, and this leads to the multiplicity of types which constitute the family of objectives. The effect of the glass constants on the basic properties of the triplet becomes therefore a very fundamental problem. It is investigated in Part II. In Part III (in preparation) the special problem of the high aperture triplet is discussed.

II. THE DESIGN OF THE TRIPLET TYPE 111 OBJECTIVE

Taylor (1893) arrived at the triplet construction by considering what happens when the components of an achromatic doublet are separated. The separation increases the power of the system without adding to the Petzval sum, and, in addition, if achromatism is to be maintained after separation, the power of the flint lens must be increased. Each of these effects provided a means for reducing the Petzval sum. The separated doublet, however, has considerable distortion and transverse chromatic aberration. Taylor met this situation by splitting the convex lens into two parts and placing these one on either side of the negative lens. From the method of its evolution, then, the triplet is to be regarded as a modified dialyte.

As already described (Cruickshank 1956) another view of the derivation of the triplet leads to more profitable procedures for its design. In 1812 Wollaston proposed an objective for the camera obscura consisting of a single meniscus positive lens placed behind a diaphragm. The triplet objective may be regarded logically as derived from the Wollaston lens by the addition of a compound correcting system comprised of a positive and a negative lens placed in front of the diaphragm. The function of this corrector system is to introduce aberrations which will compensate those of the single positive lens behind the diaphragm. In many triplets it is found that the positive and negative lenses in front of the diaphragm constitute a system of zero or very low power. Investigation shows, however, that it is not necessary to limit the corrector to zero or very low power.

(a) *The Initial Arrangement of the System*

Consider the important practical case of the initial design of a triplet photographic objective corrected for an object plane at infinity. Three different glasses will be used for the components. Disregarding the axial thickness of the lenses, five parameters are required to specify the system initially, namely, the powers φ_a , φ_b , φ_c , and the separations t_1 , t_2 , of the three co-axial thin lenses a , b , and c . The following five conditions may then be fulfilled: (i) the power of the system shall be unity; (ii) the power of the corrector system shall be χ ; (iii) the system shall have a value R_4 for the Petzval sum; and for an object plane at infinity the system shall have (iv) a residual axial chromatic aberration, R_6 , given in transverse measure, and (v) a residual transverse chromatic aberration R_7 for a pencil of obliquity u_a and a diaphragm coincident with the central lens b . Using well-known relations for systems of separated thin lenses in air these five conditions may be formulated analytically as follows:

$$1/y_{0a} \sum_{j=a}^c \varphi_j y_{0j} = 1, \quad \dots\dots\dots (1)$$

$$1/y_{0a} \sum_{j=a}^b \varphi_j y_{0j} = \chi, \quad \dots\dots\dots (2)$$

$$\sum_{j=a}^c \varphi_j / N_j = R_4, \quad \dots\dots\dots (3)$$

$$1/u'_{0c} \sum_{j=a}^c \varphi_j y_{0j}^2 / V_j = R_6, \quad \dots\dots\dots (4)$$

$$1/u'_{0c} \sum_{j=a}^c \varphi_j y_{0j} y_j / V_j = R_7, \quad \dots\dots\dots (5)$$

where y_j and y_{0j} are the incidence heights at the j th component of a principal paraxial ray of obliquity u_a and an axial paraxial ray respectively, and u'_{0j} is the inclination angle of the axial ray after refraction at the j th component.

Since the diaphragm is initially in coincidence with the second thin lens of the system the incidence heights of the principal paraxial ray are such that

$$y_a/y_c = -t_1/t_2, \quad y_b = 0, \quad \dots\dots\dots (6)$$

while for the axial ray

$$\begin{aligned} y_{0b} &= y_{0a} - t_1 u'_{0a} \\ &= y_{0a} (1 - t_1 \varphi_a), \quad \dots\dots\dots (7) \end{aligned}$$

and

$$y_{0c} = y_{0b} - t_2 y_{0a} \chi. \quad \dots\dots\dots (8)$$

A substantial reduction in symbols specifying the glasses of the system is achieved by writing

$$V_a/V_b = \alpha, \quad V_a/V_c = \xi, \quad \dots\dots\dots (9)$$

$$N_a/N_b = \beta, \quad N_a/N_c = \gamma. \quad \dots\dots\dots (10)$$

In addition we will also write

$$y_{0b}/y_{0a} = \eta_{0b}, \quad y_{0c}/y_{0a} = \eta_{0c}, \quad \dots\dots\dots (11)$$

$$R_4 N_a = P, \quad R_6 V_a / y_{0a} = L, \quad R_7 V_a / u_a = T \quad \dots\dots\dots (12)$$

With this notation equations (1)–(8) now become

$$\varphi_a + \eta_{0b}\varphi_b + \eta_{0c}\varphi_c = 1, \quad \dots\dots\dots (13)$$

$$\varphi_a + \eta_{0b}\varphi_b = \chi, \quad \dots\dots\dots (14)$$

$$\varphi_a + \beta\varphi_b + \gamma\varphi_c = P, \quad \dots\dots\dots (15)$$

$$\varphi_a + \alpha\eta_{0b}^2\varphi_b + \xi\eta_{0c}^2\varphi_c = L, \quad \dots\dots\dots (16)$$

$$(1+T)t_1\varphi_a - t_2\xi\eta_{0c}\varphi_c = T, \quad \dots\dots\dots (17)$$

$$1 - t_1\varphi_a = \eta_{0b}, \quad \dots\dots\dots (18)$$

$$\eta_{0b} - t_2\chi = \eta_{0c}. \quad \dots\dots\dots (19)$$

The solution of these equations is quite simple. Equations (13) and (14) give at once

$$\eta_{0c}\varphi_c = 1 - \chi, \quad \dots\dots\dots (20)$$

and, combining this with equations (17)–(19), we obtain

$$t_1 = (1 - \eta_{0b})/\varphi_a, \quad \dots\dots\dots (21)$$

$$t_2 = [1 - (1+T)\eta_{0b}]/\xi(1-\chi). \quad \dots\dots\dots (22)$$

Substitution in (19) gives

$$\kappa\eta_{0b} - \xi(1-\chi)\eta_{0c} = \chi, \quad \dots\dots\dots (23)$$

where

$$\kappa = \xi + \chi(1 - \xi + T). \quad \dots\dots\dots (24)$$

Subtracting (14) from (16) yields

$$\varphi_b\eta_{0b}(\alpha\eta_{0b} - 1) + \xi(1-\chi)\eta_{0c} = L - \chi,$$

which on combination with (23) leads to

$$\varphi_b = (\kappa\eta_{0b} - L)/(1 - \alpha\eta_{0b})\eta_{0b}. \quad \dots\dots\dots (25)$$

Equations (20) and (21) give at once

$$\varphi_c = \xi(1-\chi)^2/(\kappa\eta_{0b} - \chi), \quad \dots\dots\dots (26)$$

while (14) and (25) give

$$\begin{aligned} \varphi_a &= \chi - \eta_{0b}\varphi_b \\ &= [L + \chi - (\kappa + \alpha\chi)\eta_{0b}]/(1 - \alpha\eta_{0b}). \quad \dots\dots\dots (27) \end{aligned}$$

Eliminating φ_a , φ_b , φ_c from equation (15) by means of (25)–(27) we obtain the cubic equation

$$G_3\eta_{0b}^3 + G_2\eta_{0b}^2 + G_1\eta_{0b} + G_0 = 0, \quad \dots\dots\dots (28)$$

where

$$G_3 = \kappa^2 + \alpha\kappa(\chi - P), \quad \dots\dots\dots (29)$$

$$G_2 = \kappa(P - L - 2\chi - \beta\kappa) + \alpha[\gamma\xi(1-\chi)^2 - \chi(\chi - P)], \quad \dots\dots (30)$$

$$G_1 = \beta\kappa(\chi + L) - [\gamma\xi(1-\chi)^2 - \chi(\chi - P)] + L\chi, \quad \dots\dots\dots (31)$$

$$G_0 = -L\beta\chi. \quad \dots\dots\dots (32)$$

These coefficients are computed and the cubic equation (28) is easily solved either by, say, Cardan's method or by successive approximation. In general only one root of this equation gives a physically useful solution. There is an important exception, however, in one case in which a second real root leads to the possibility of the construction of another group of objectives of high aperture. These will be considered separately later. With the value of η_{0b} obtained from (28), the values of the powers and separations are calculated from the appropriate foregoing equations, and the initial arrangement is thus determined. If $L=0$, the coefficient G_0 vanishes and the cubic (28) reduces to a quadratic.

(b) *Triplets from Two Glasses Only*

It is very common practice to use two glasses only in the construction of the triplet, the two positive components a and c being of the same crown glass. In this case, then,

$$\gamma=1=\xi, \quad \kappa=1+\chi T.$$

Inserting these values in the equations of the previous section we obtain in place of (28) the modified cubic

$$g_3\eta_{0b}^3 + g_2\eta_{0b}^2 + g_1\eta_{0b} + g_0 = 0, \quad \dots\dots\dots (33)$$

in which

$$g_3 = \kappa^2 + \alpha\kappa(\chi - P), \quad \dots\dots\dots (34)$$

$$g_2 = \kappa(P - L - 2\chi - \beta\kappa) + \alpha[1 + \chi(P - 2)], \quad \dots\dots\dots (35)$$

$$g_1 = \beta\kappa(\chi + L) + \chi(L - P + 2) - 1, \quad \dots\dots\dots (36)$$

$$g_0 = -L\beta\chi. \quad \dots\dots\dots (37)$$

The expressions for the powers and separations become

$$\varphi_b = (\kappa\eta_{0b} - L)/(1 - \alpha\eta_{0b})\eta_{0b}, \quad \dots\dots\dots (38)$$

$$\varphi_c = (1 - \chi)^2/(\kappa\eta_{0b} - \chi), \quad \dots\dots\dots (39)$$

$$\varphi_a = \chi - \eta_{0b}\varphi_b, \quad \dots\dots\dots (40)$$

$$t_1 = (1 - \eta_{0b})/\varphi_a, \quad \dots\dots\dots (41)$$

$$t_2 = [1 - \eta_{0b}(1 + T)]/(1 - \chi). \quad \dots\dots\dots (42)$$

(c) *Discussion of the Solution for a Given Set of Glasses*

The general characteristics of the equations of the previous sections should now be considered. Suppose that a selection of three glasses is made for the system and a set of values chosen for the aberration residuals P , L , and T . The G coefficients in equation (28) depend thereafter only on the value of χ , so that in this situation η_{0b} , φ_a , φ_b , φ_c , t_1 , and t_2 are effectively functions of χ only. It is necessary to consider then how these quantities vary with χ , i.e. how the initial arrangement of the triplet depends on the power of the corrector system. In Figures 2 and 3 the variation of each parameter with χ is shown for a typical triplet in which

$$\begin{array}{lll} N_a = 1.6226 & V_a = 60.2 & P = 0.40 \\ N_b = 1.61706 & V_b = 36.53 & L = 0 \\ N_c = 1.6226 & V_c = 60.2 & T = 0 \end{array}$$

It is clear that values of χ in the range $0.4 < \chi < 1$ are to be excluded because the discontinuity which occurs here either provides no solution or one in which the powers of the components and the large rear airspace are totally unsuitable. For values of $\chi > 1$, φ_c and t_2 become negative. The system would then consist of two negative lenses with a strong positive lens between them, but the negative separation renders the arrangement impracticable, at least with simple lenses.

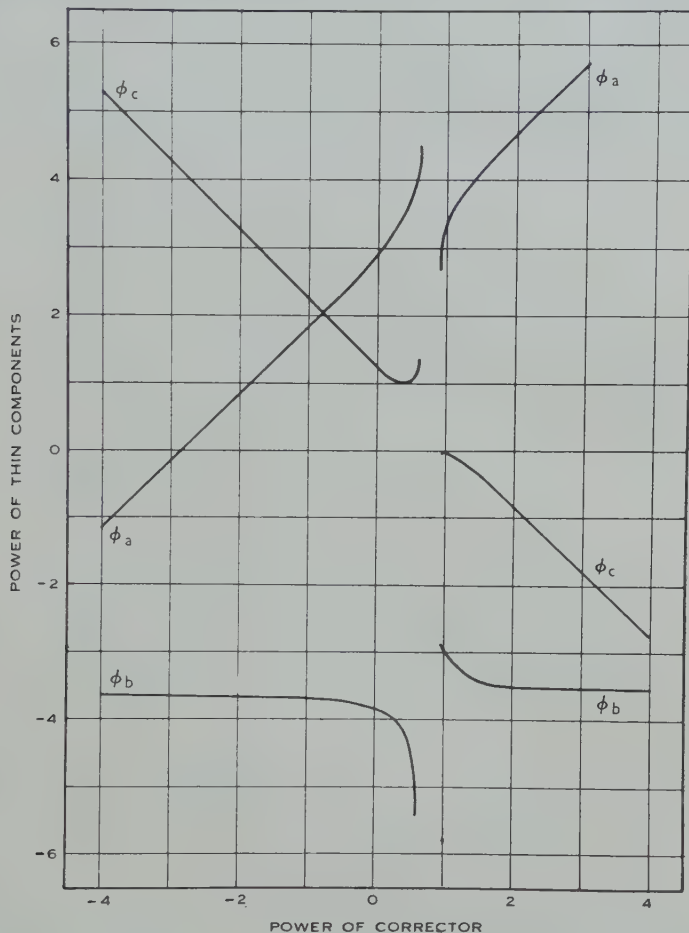


Fig. 2.—The curves show the dependence of the powers, φ_a , φ_b , φ_c , of the three thin components of a unit power triplet upon the parameter χ . The useful range of χ -values is seen to be approximately $-2 < \chi < 0.4$.

Again, if $\chi < -2.2$, the rapidly increasing front airspace renders the solution useless, in addition to which φ_a and t_1 soon become negative, providing again an impracticable arrangement. We are left then with the values $-2 < \chi < 0.4$ as approximately defining the useful range of χ .

It will be noted that in this range φ_b changes very little with χ , while φ_a and φ_c vary rapidly and almost linearly with χ . Moreover, the slopes of these

two lines are almost opposite. Within this practical range of χ -values, therefore, a change in χ leaves the negative lens almost unaffected but results in an exchange of power between the two positive lenses. In addition, as Figure 3 shows, there are changes in the airspaces. At one end of the range the front airspace is small and the back airspace large, while at the other end of the range the situation is reversed. Near the middle of the range the total thickness of the system is a minimum. *The essential point is that χ determines the distribution of power between the two positive components of the system.*

There now remains the consideration of the effect on the solutions of variations of the aberration residuals, P , L , and T . Typical changes in the powers and separations due to the variation of P alone are shown in Figures 4

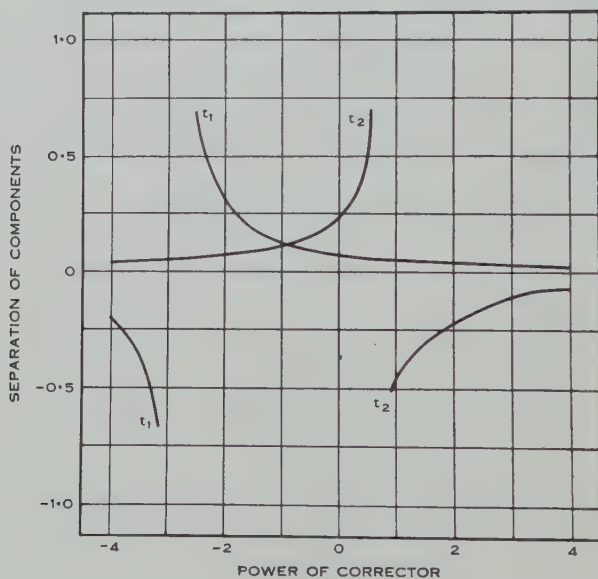


Fig. 3.—The curves show how the airspaces, t_1 and t_2 , vary with the parameter χ in a unit power triplet employing the glasses specified in Section II (c).

and 5. As is to be expected from equation (15), an increasing positive value of P is accompanied by reduction in the power of all components, though the power of the back lens is affected least. This reduction in the curvatures throughout the system following relaxation of the Petzval condition is accompanied by increase of the front airspace and decrease of the rear airspace. In Figures 6 and 7 the effects of variation of the residual L are shown. If it is sought to adjust the longitudinal chromatic aberration of an objective to zero for the 0.7 zone of the aperture, as is frequently desired, an appropriate positive residual value for L will be required. The figures show that this will result in reductions in the powers of the first two lenses and a very slight increase in the power of the third lens. The accompanying increases in the airspace are quite substantial, however, and have to be considered carefully in the design of the triplet. There is little

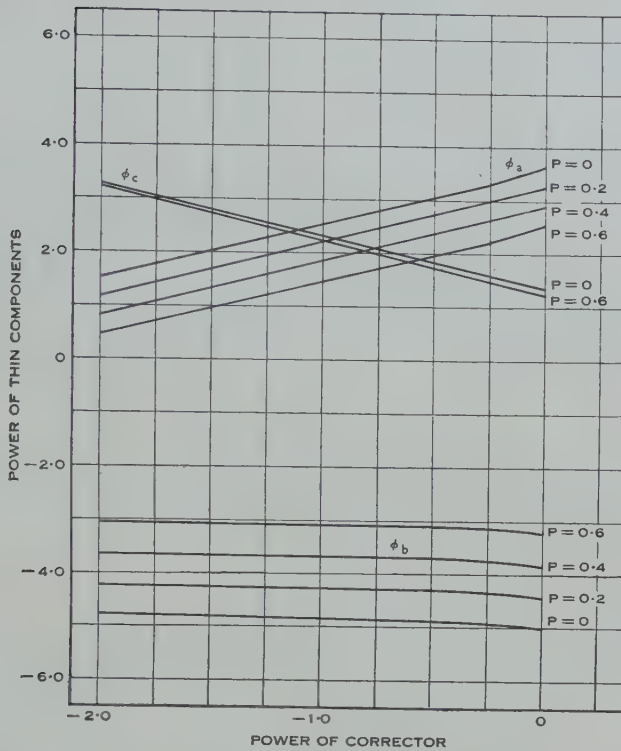


Fig. 4.—The graphs show the effect of the variation of the Petzval residual P upon the powers, ϕ_a , ϕ_b , ϕ_c , of the thin components of a unit power triplet. Increase of P reduces the absolute values of the powers of lenses a and b , but has little effect upon that of lens c .

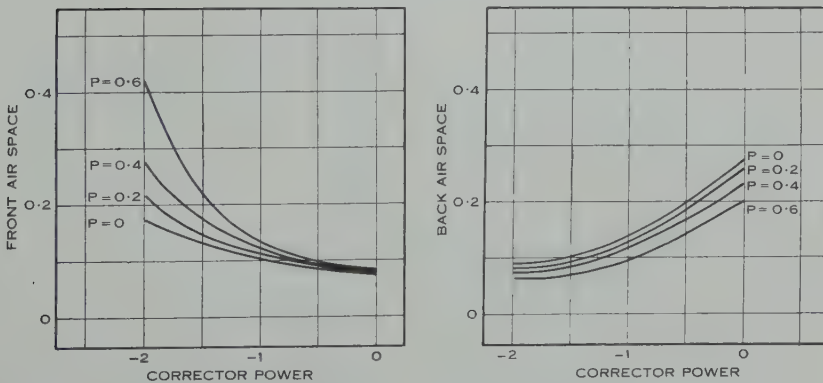


Fig. 5.—The corresponding effect upon the airspaces accompanying variation of the Petzval residual P is shown. An increase in the value of P enlarges the front airspace, particularly for large negative values of χ , and diminishes the back airspace. For values of χ near zero it is the rear airspace that is mainly affected.

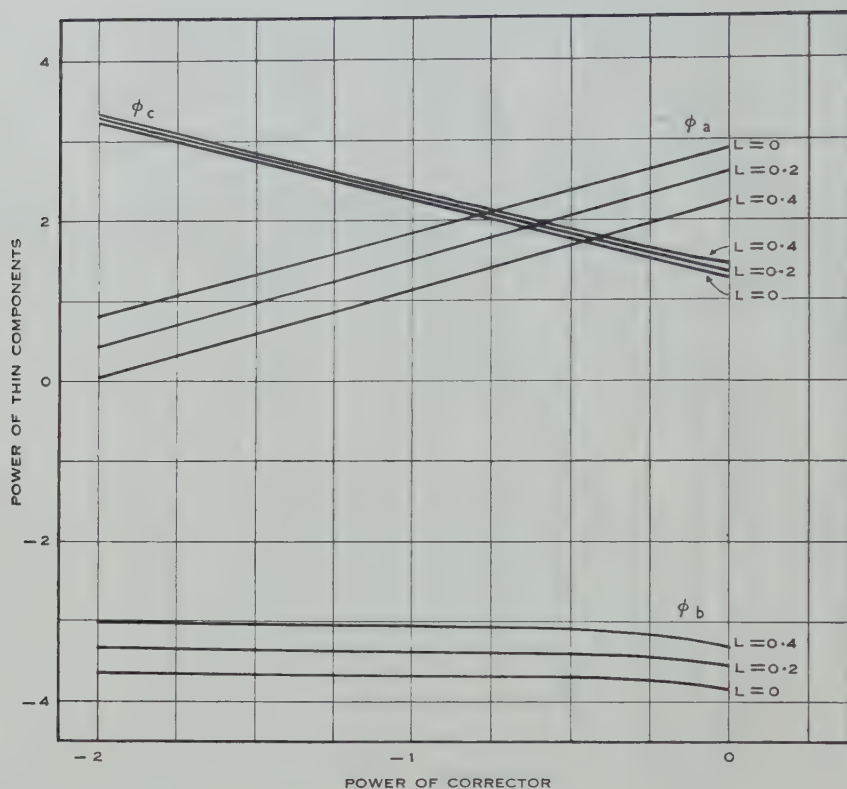


Fig. 6.—The effect of the variation of the axial chromatic aberration residual L upon the powers of the thin components is shown. A positive value of L , corresponding to a correction of longitudinal colour at an outer zone of the aperture, reduces the total curvatures in the first two components, but increases the power of the rear positive lens slightly.

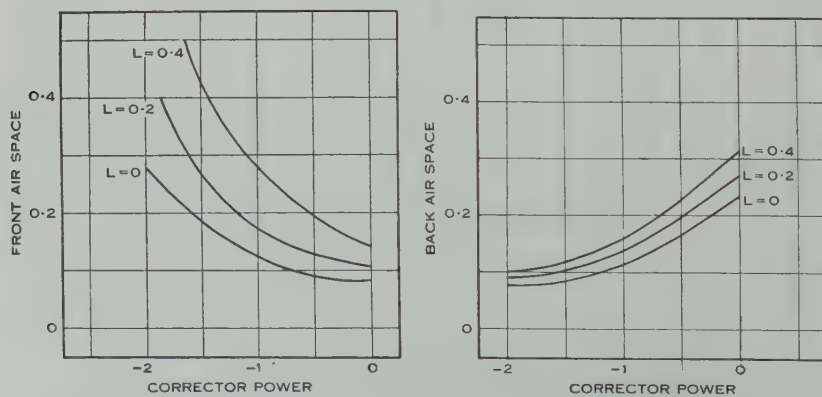


Fig. 7.—The effect on the airspaces of a variation of L is exhibited in these graphs. Both separations are enlarged by the introduction of a positive residual value for L .

effect from the variation of T over the small range of values which may be required to achieve a satisfactory compromise for the transverse chromatic aberration across the field. The changes in the powers and separations of the components are only of the order of 1 or 2 per cent. of their values and have not been plotted. This statement as to T may require modification later when consideration is given to the design of objectives with compound members.

(d) *The Adjustment of Primary Coma, Astigmatism, and Distortion*

For the adjustment of the primary astigmatism, distortion, and coma we use the degrees of freedom available in the selection of the shapes, S_a , S_b , S_c , of the three components. Either the shape function introduced by Coddington or that of Argentieri is suitable. Using the Coddington shape function defined by

$$S = (c_1 + c_2)/(c_1 - c_2),$$

where c_1 and c_2 are the curvatures of the two surfaces of the thin lens, it is well known that for a system of separated thin lenses the coefficients of the primary coma, astigmatism, and distortion are given respectively by equations of the form

$$\sigma_2 = \Sigma(a_{21}S^2 + a_{22}S + a_{23}), \quad \dots\dots\dots (43)$$

$$\sigma_3 = \Sigma(a_{31}S^2 + a_{32}S + a_{33}), \quad \dots\dots\dots (44)$$

$$\sigma_5 = \Sigma(a_{51}S^2 + a_{52}S + a_{53}). \quad \dots\dots\dots (45)$$

The partial coefficient of each lens in respect of each of these aberrations is thus a quadratic function of its shape and the summation is extended over all the lenses. If the diaphragm coincides with one of the thin lenses, then for that lens

$$a_{21} = a_{31} = a_{32} = a_{51} = a_{52} = a_{53} = 0.$$

In the foregoing work the diaphragm of the triplet has been set initially in coincidence with lens b . If the desired values of σ_3 and σ_5 are set at zero or some other small residual, equations (44) and (45) provide a pair of simultaneous equations, quadratic in S_a and S_c , which can be solved for the shapes of these lenses. Substitution of the values so obtained in (43) reduces this equation to one linear in S_b , which may then be solved to give zero or any other desired value of the coma. This provides a straightforward means of determining the shapes of the lenses for any triplet arrangement which will control the primary astigmatism, distortion, and coma.

(e) *Adjustment of the Spherical Aberration*

There remains for consideration, finally, the adjustment of the primary spherical aberration. Taylor (1904) pointed out long ago that the means for this lies in the distribution of the total positive power between the two collective lenses. An essential feature of the present treatment is that a variable χ has been introduced at the outset which (i) permits the initial solution to be easily made and (ii) specifies the distribution of power between the collective lenses. It is therefore the variable which controls the primary spherical aberration. It is not easy to derive a useful exact expression for the spherical aberration as

a function of χ . Some elementary approximate reasoning, however, leads to the result that within the range of values of practical interest we should have

$$\sigma_1 \simeq a_0 + a_1\chi + a_2\chi^2 + a_3\chi^3,$$

where the a 's are constant and a_3 generally small. It might be expected then that the spherical aberration could be closely approximated by a quadratic function of χ . Investigation shows this to be the case, the actual curve being adequately fitted by a quadratic function of χ . The coefficients of the quadratic expression are then determined from the computed spherical aberration of three systems with different corrector powers.

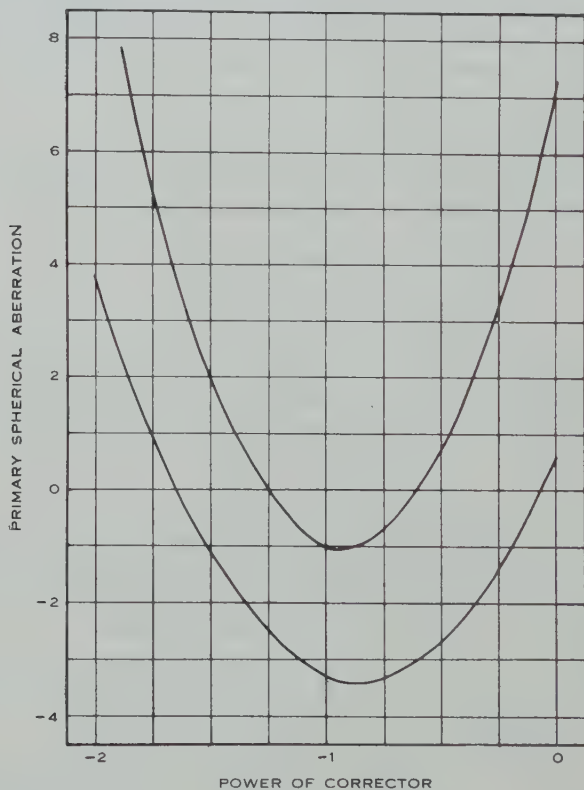


Fig. 8.—Curves showing the variation of the primary spherical aberration coefficient σ_1 with the parameter χ in unit power triplets from a given set of glasses. All other primary aberrations have been reduced to selected small values. The upper curve is for thin lenses and the lower curve for triplets with components of finite thicknesses.

In Figure 8 the upper curve shows the variation of the primary spherical aberration with χ in triplets from a given selection of glass and prescribed values of P , L , and T . All triplets represented in the curve are corrected for all other primary aberrations. As the curve intersects the axis in two points there are two values of χ for which the primary spherical aberration has some prescribed small value. Thus there are two possible types of solutions corresponding to

quite different values of χ . Triplets quoted in the literature belong principally to the right-hand branch of the parabola, i.e. they belong to the type having the less negative value of χ . The other type with the more negative value of χ deserves to be investigated more because for certain purposes it is to be preferred.

(f) *The Thickening of the System*

It is now necessary to replace each thin component by one of finite axial thickness sufficient to provide the desired aperture for the component and the necessary edge thickness for easy manufacture. This process results, of course, in the introduction of residuals for all primary aberrations and these must be eliminated. A satisfactory procedure is to take the thin component system which has been computed and replace the thin components by lenses of appropriate axial thickness and compute the new primary aberrations. The changes in chromatic aberration and Petzval curvature are usually negligible. Returning to equations (44), (45), the derivatives $\partial\sigma_3/\partial S_a$, $\partial\sigma_3/\partial S_c$, $\partial\sigma_5/\partial S_a$, and $\partial\sigma_5/\partial S_c$ are calculated and hence the changes in the shapes, S_a and S_c , of the thin system necessary to give astigmatism and distortion residuals equal in amount, say, but opposite in sign to those introduced by the previous thickening. The system with the new shapes is then thickened and the new primary aberrations determined. Graphs may then be drawn of the astigmatism and distortion of the thin system against the astigmatism and distortion respectively of the thickened system, assuming the relations to be linear. The thin system residual corresponding to the desired thick system residual is read off from the graph and the final shapes computed. Similarly the coma of the system is adjusted by changing slightly the value of S_b . When this system is thickened it will be found to have primary aberration values very close to those desired.

In this way three triplets of different corrector powers are obtained having thickened components and all primary aberrations controlled except spherical aberration. The lower curve in Figure 8 represents the variation of spherical aberration with χ in thickened triplets in which all other primary aberrations have been adjusted to desired small values. This curve is fitted by a quadratic function of χ and so can be drawn from three calculated points. From it may be read the value of χ necessary to achieve any desired residual of spherical aberration. Normally a small positive residual of spherical aberration is required to offset the negative secondary and tertiary terms. The two values of χ which will provide these residuals are determined and the solutions of these triplets developed.

To complete the design it is necessary to obtain a good balance between the primary and higher order aberrations. Using the new and very effective means introduced by Buchdahl (1954) the secondary aberrations may then be calculated and the balance inspected. In a particular case it is generally a fairly straightforward matter to see how the balance of aberrations could be improved, and the primary design is accordingly altered to secure this.

(g) *Finite Object Distance*

In the cases considered so far the object has been confined to a plane at infinity. If this restriction is removed some of the basic equations of Section (b)

become more complex and the writer has only obtained a general solution of these for the special case for which $\chi=0=L=T$.

There is, however, another simple way to deal with this problem. Of the five initial conditions the first three are independent of the position of the object, while the remaining two conditions, those for achromatism, do involve the object position. The problem may therefore be solved initially for an infinitely distant object position and the chosen values of L and T . The change in the chromatic aberration of the thin component system may then be calculated when the object plane is moved to the desired position at a finite distance from the lens. The initial solution may then be repeated with the values of L and T altered by amounts depending on the chromatic changes introduced by the shift of the object plane.

(h) Recapitulation

The method of design just described may be summarized as follows :

- (i) A set of glasses is chosen and residual values of the Petzval sum and the chromatic aberrations are prescribed.
- (ii) With these constants the initial equations are solved for three triplets with different values of χ , say $\chi=0$, -0.5 , -1.0 .
- (iii) The shapes are determined in each of these so that astigmatism, distortion, and coma are controlled as desired.
- (iv) The systems are thickened and the shapes readjusted to maintain correction of the three oblique aberrations.
- (v) The curve of primary spherical aberration against χ is plotted and the values of χ selected which will give the desired primary spherical aberration.
- (vi) Either or both of these solutions are then developed.

It may be thought that this is a very long procedure. What is done, of course, is to survey the complete possibilities with one set of glasses and residuals, which is more than the development of the design of one triplet. At every stage there is complete control of all the factors affecting the design and a clear understanding of the effect of the variation of each parameter on the primary aberrations. After the method has been used a little it will be realized that there are a number of short cuts which can be taken due to the experience acquired, which reduce the work considerably.

III. ACKNOWLEDGMENTS

The author wishes to acknowledge gratefully the help given by Mrs. B. J. Brown, who has been responsible for the whole of the computational work associated with the development of satisfactory procedures in the different phases of the methods described. Thanks are due also to the Department of Defence Production and to the firm of E. N. Waterworth for financial support of this work.

IV. REFERENCES

- BUCHDAHL, H. A. (1954).—"Optical Aberration Coefficients." (Oxford Univ. Press.)
 CRUICKSHANK, F. D. (1956).—*Rev. Opt. (Théor. Instrum.)* **35**: 292.
 TAYLOR, H. D. (1893).—Brit. Pat. 22607/93.
 TAYLOR, H. D. (1904).—*Mon. Not. R. Astr. Soc.* **1904**: 615.

FURTHER OBSERVATIONS OF RADIO EMISSION FROM THE PLANET JUPITER

By F. F. GARDNER* and C. A. SHAIN*

[Manuscript received July 24, 1957]

Summary

The results of further observation of Jupiter radiation, made near Sydney from June 1955 to March 1956, are presented. Most work was done at 19.6 Mc/s, but some observations were also made at 14 and 27 Mc/s.

The characteristics of the 19.6 Mc/s radiation received from Jupiter are described in detail. The new facts disclosed by these observations are: (i) Jupiter radiation appeared to be random noise varying rapidly in intensity—large changes in intensity took place in times as short as 0.2 sec, but no shorter; (ii) some spaced-receiver observations indicated that the terrestrial ionosphere can have a pronounced effect on the short-term characteristics of the radiation; (iii) a single observation showed that on that occasion the radiation was circularly polarized (right-handed in radio sense).

There appeared to be three sources of noise on Jupiter. None of these could be identified with visual features. The radiation emitted from the main source was confined to angles within 45° of a central line.

The additional observations at 14 and 27 Mc/s showed that the peak intensity, mean duration, and frequency of occurrence were all greatest at 19.6 Mc/s. Two of the radio sources were active at 19.6 Mc/s but not at 27 Mc/s. The angular spread of the radiation from the principal source was considerably narrower at 27 Mc/s than at 19.6 Mc/s.

The great variability and spectral concentration of the Jupiter radiation, both resembling solar noise, suggest an origin in some form of plasma oscillation in an ionized region with a critical frequency around 20 Mc/s.

I. INTRODUCTION

Radio noise from Jupiter was first observed early in 1955 by Burke and Franklin (1955) at Washington, with a 22.3 Mc/s "Mills Cross" system. The intensity of the radiation was very great at 22.3 Mc/s, but appeared to fall off rapidly with increasing frequency, as no radiation was observed at all at 38 Mc/s. After the discovery was announced, a search by one of us (Shain 1955, 1956) of old records of cosmic noise taken in Sydney at 18.3 Mc/s confirmed the American observations and showed in addition that in 1951 the radiation came from a localized region on the planet. The absence of high frequency radiation was confirmed by an examination of records taken at Cambridge at 38 Mc/s (Smith 1955) and at Sydney on 85 Mc/s.

At this stage our knowledge of the basic facts concerning the radiation was too incomplete for the drawing of worth-while conclusions about the mechanism of generation of the radiation. The further observations to be described were

* Division of Radiophysics, C.S.I.R.O., University Grounds, Chippendale, N.S.W.

an attempt to answer some of the outstanding problems. These included the spectrum and polarization of the radiation, the location and features of the planetary source, and particularly the characteristic time variations of the noise itself. The results obtained have provided only partial answers to these problems. They suggest an origin of the noise in some type of resonance oscillation in Jupiter's atmosphere, but optical data give no clues as to where or why such oscillations occur.

II. EQUIPMENT

The observations were made at Fleurs (34°S., 151°E.), near Sydney, from June 1955 to March 1956. During this time Jupiter's declination varied between +15° and +20°, corresponding to zenith angles at transit of 49° and 54°. Transit was at midday on August 5, 1955 and at midnight on February 19, 1956.

Because the radiation from Jupiter is very variable and bears some resemblance to interfering signals from terrestrial atmospherics, solar noise, and radio stations, which can be very severe at frequencies near 20 Mc/s, it was

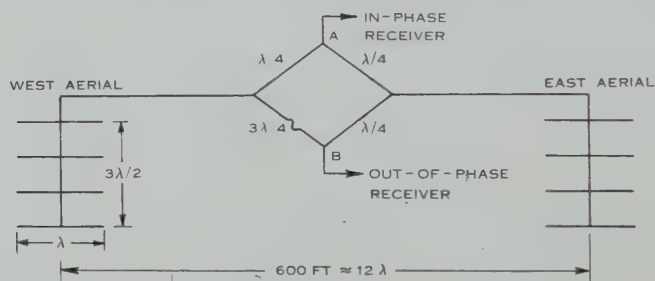


Fig. 1.—19.6 Mc/s interferometer receiving system.

considered essential to have a good system for identifying the radiation from the planet. Accordingly the main equipment comprised the 19.6 Mc/s interferometer system shown diagrammatically in Figure 1. Each aerial of the interferometer consisted of four full-wave dipoles, each a quarter-wavelength above ground, and phased to produce maximum response to the north at a zenith distance of about 50°. Since the aerial was only moderately directive in azimuth, Jupiter could be observed over a period of about 5 hr per day. The two aerials were spaced about 12 wavelengths apart in an east-west direction, so that the angle between minima of the interference pattern was about 5°. The bridge system shown in Figure 1 permitted the connexion of the aerials to two receivers which were effectively isolated from one another, one using the aerials in-phase, the other out-of-phase. Figure 2 shows a sample record obtained with this system, and also the response expected for a steady source on Jupiter. The radiation is much stronger on the "in-phase" receiver near the transit of Jupiter (at the time marked A in the figure) and near C, 20 min earlier; near the intermediate point B it is stronger in the "out-of-phase" receiver. The system has the advantage that even short duration bursts can be identified as Jupiter radiation by the ratio of the outputs of the two receivers.

At times, observations were made with the east-west arm of the 19.7 Mc/s Cross aerial at Fleurs, a 3400 ft in-line array of dipoles one quarter-wavelength above ground, and also with a small aerial at Potts Hill, 25 km east of Fleurs.

For obtaining information on the spectrum of the radiation, single in-line arrays of four and eight half-wave dipoles were used for reception on frequencies of 14 and 27 Mc/s respectively. The dipoles were fed in-phase and their height above ground was such as to give maximum response 50° from the vertical and in the north-south plane.

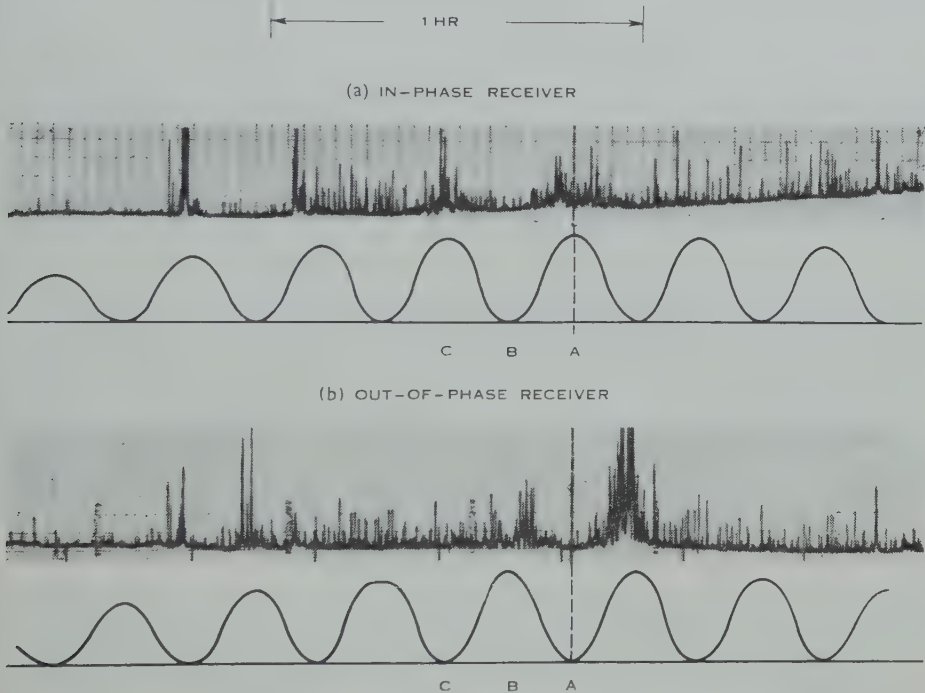


Fig. 2.—Typical record obtained with the 19.6 Mc/s interferometer system, June 28, 1955. Below each record the responses of the two receivers to a steady source on Jupiter are shown. The fine spikes on the record are atmospherics; all the remaining increases are Jupiter radiation.

To measure polarization, the east-west arm of the Cross was used as an interferometer in conjunction with a small array whose individual dipoles were in the north-south plane and also perpendicular to the direction of Jupiter at transit. Because of the narrow beam of the long east-west array, polarization measurements were confined to times within 10 min of transit. The ratio, in amplitude and phase, of the two transverse components of the downcoming wave was measured. The method used was more accurate for phase than for amplitude.

III. RESULTS

(a) Time Variation of the Radiation

The characteristics of the radiation received from Jupiter are illustrated by the records of Figure 3 which were made at three different chart speeds. In the

upper pair of records of Figure 3 (a), similar to those of Figure 2, the radiation from Jupiter lasts about 40 min, from about 09^h 20^m to 10^h (sidereal time) on December 26, 1955. The radiation did not occur every day, but on days when it did occur the average overall duration of the activity was about half an hour out of the possible observing time of 5 hr.

A distinctive feature of both Figures 2 and 3 (a) is the occurrence of groups of bursts of radiation lasting for times of the order of a minute. They seem to be characteristic of the emission and not produced by the terrestrial ionosphere (see below). In the lower pair of records of Figure 3 (a), where the bracketed

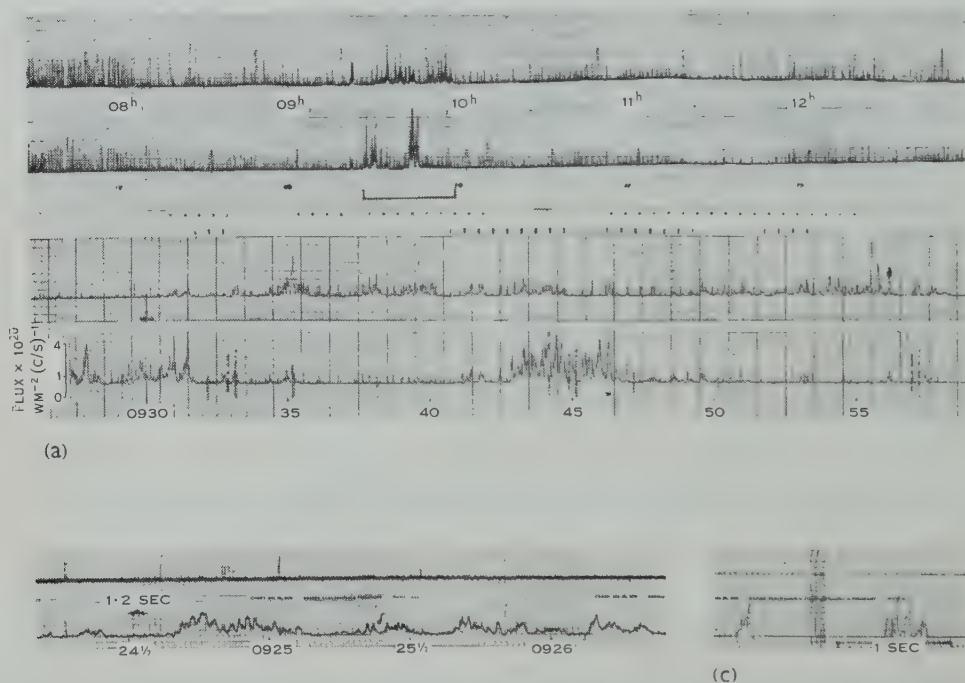


Fig. 3.—Interferometer records taken with two receivers at 19.6 Mc/s. All times given are local sidereal time for Fleurs (longitude 150° 46.4' E.). (a) The usual type of record; the upper pair was taken at low paper speed, 6 in/hr, and the bracketed portion is shown below at 10 times this speed, 1 in/min, December 28, 1955. (b) High speed record of Jupiter taken with Brush Recorder at 20 in/min, near the time of minimum response of the in-phase receiver, November 1, 1955. (c) Atmospherics recorded at highest speed, 5 in/sec, one receiver only in use.

portion of the upper pair is shown with 10 times the chart speed, it can be seen that fluctuations within each burst are very pronounced. Frequent breaks of a minute or so occur, during which there is no radiation from Jupiter. The numerous short spikes on the records are atmospherics—identified as such because their envelope does not follow the characteristic interference pattern of the aerial.

To study the faster fluctuations which are not resolved in Figure 3 (a), a high-speed mechanical oscillograph* with a response time of a few milliseconds

* Brush Electronics Company Type BL 202, kindly lent by Mr. G. Reber.

was employed. A typical pair of records is shown in Figure 3 (b), from which it would appear that there are no significant bursts much shorter than 1 sec, although rise times of some bursts may be as short as one-fifth of a second; any changes faster than about $\frac{1}{5}$ sec are only of the size expected for random noise of the appropriate intensity. The shorter spikes are attributed to atmospherics, since their relative amplitudes did not follow the interference pattern of the aerial, which is near a maximum on the lower trace and a zero on the upper. The atmospherics provide a check on the response time of the system. Figure 3 (c) shows an atmospheric recorded at 25 times the chart speed of Figure 3 (b). It is seen that the receiver-recorder system can respond to sudden changes in about 10 msec, and yet on no occasion has there been any evidence for Jupiter bursts of the order of 10 msec, as reported by Kraus (1956).

When heard on a loudspeaker the Jupiter noise sounded like thermal noise varying rapidly in intensity, but only at a rate which the ear could follow. The overall impression of the noise was a resemblance to solar noise rather than terrestrial atmospherics. The time structure appeared similar to that of enhanced solar radiation at about 100 Mc/s.

An intensity scale is shown on the lowest record of Figure 3 (a). On this occasion the intensity was higher than usual, the peaks exceeding $4 \times 10^{-20} \text{ W m}^{-2} (\text{c/s})^{-1}$. The highest peak intensity recorded was about $10^{-19} \text{ W m}^{-2} (\text{c/s})^{-1}$, which is of the same order as that of a small solar outburst and more than 100 times greater than that of the strongest radio star at 20 Mc/s. On different occasions the intensity ranged from the maximum just quoted down to the minimum detectable level of about $10^{-21} \text{ W m}^{-2} (\text{c/s})^{-1}$.

(b) *Spaced-receiver Experiment*

The angular size of the source on Jupiter is very much smaller than that of any other known radio source, and it would be expected that very severe fluctuations in intensity would sometimes be caused by scintillations in the Earth's ionosphere. In an attempt to answer the question of how important are these scintillations, a short series of simultaneous observations were made at two sites (Fleurs and Potts Hill) separated by 25 km in an east-west direction. As a precaution against spectral variations in intensity, the receivers at the two sites were tuned as closely as possible (within about 2 kc/s) to the same frequency, using a portable crystal oscillator as reference. Unfortunately there was considerable local interference at Potts Hill and the number of pairs of records available for comparison is only 3. These were taken between 23 and 02 hr local time. However, these few records are sufficient to show that there were considerable differences in the time variations of intensity at the two sites.

A striking example is shown in Figure 4, which shows the records obtained over about a quarter of an hour near 10^h sidereal time (about 23^h 30^m Eastern Australian Standard Time) on February 26, 1956. The top record is from Potts Hill, and the bottom record is made up of complementary sections of the two interferometer records from Fleurs. It will be seen that, although there is a general similarity between the Fleurs and Potts Hill records, there are very marked differences. For example, the bursts recorded at Fleurs between

the ground was coherent in phase and amplitude over a distance of about 200 m. On the other hand, the present spaced-receiver experiment shows that the size of the irregularities was probably less than 25 km. It would be interesting to find out whether these irregularities are the same as those (with dimensions of about 4 km) which are studied in observations of discrete source scintillations (see, for example, Hewish 1951).

(c) *Frequency of Occurrence of the Radiation*

The observations were practically continuous from mid July 1955 until the end of March 1956. The histogram of Figure 5 (a) gives the number of days in each month on which Jupiter radiation was definitely received with the interferometer. There was a maximum of recorded activity in November when radiation was identified on about one day in three.

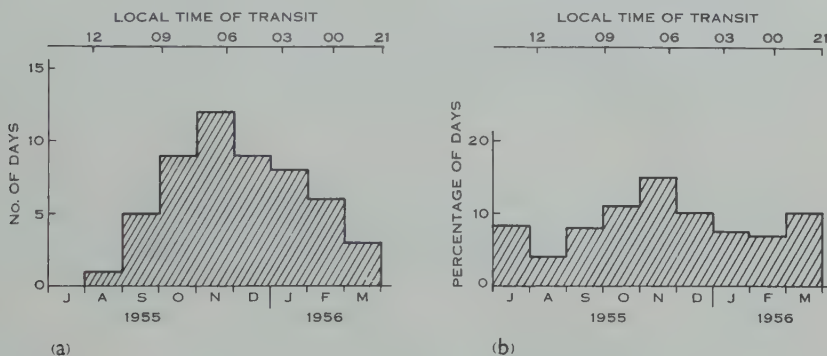


Fig. 5.—Monthly frequency of detection of Jupiter radiation on 19.6 Mc/s, (a) with the interferometer receiving system, and (b) with the receiving system using the east-west array of the Cross aerial.

The apparent variation in Jupiter activity shown in Figure 5 (a) is likely to be affected by the masking effect of interference, the average level of which varies greatly with time of day. If this is so, the observed frequency of occurrence would be expected to depend on the local time of transit, with a maximum during the early morning hours when interference is generally low. Figure 5 (a) does suggest such an effect. Observations with the east-west array of the Cross are less susceptible to interference, and the variation in the apparent frequency of occurrence of Jupiter radiation observed with this equipment, shown in Figure 5 (b), is considerably less than in Figure 5 (a). There may, however, be a true maximum in Jupiter activity in November.

(d) *Spectral Observations*

Besides the observations on 19.6 Mc/s, additional observations were made, between November and March, on frequencies of 27 and 14 Mc/s. On any day only one of the additional frequencies was used, together with the comparison frequency of 19.6 Mc/s. The information obtained on 27 Mc/s is more complete and will be considered first.

Between November 1955 and February 1956, when Jupiter was transiting between midnight and sunrise, interference on 27 Mc/s was negligible (ionospheric critical frequencies were low and long-distance propagation consequently poor) and there was no trouble from false identifications. The 27 Mc/s equipment was operating on 52 days, and the number of occasions on which Jupiter radiation was observed on these days is set out in Table 1.

TABLE 1
COMPARISON OF SIMULTANEOUS OBSERVATIONS ON 19.6 AND 27 Mc/s

Frequency on which Jupiter Radiation was Observed	Number of Days
19.6 Mc/s but not 27 Mc/s	10
27 Mc/s together with 19.6 Mc/s	10
27 Mc/s but not 19.6 Mc/s	0
No radiation at either frequency	32
Total number of days on which the 27 Mc/s equipment was operating	52

The remarkable feature of Table 1 is that on no occasion was radiation observed on 27 Mc/s but not on 19.6 Mc/s, even although all increases in noise on 27 Mc/s when Jupiter was in the aerial beam, other than atmospherics, were classed as Jovian.

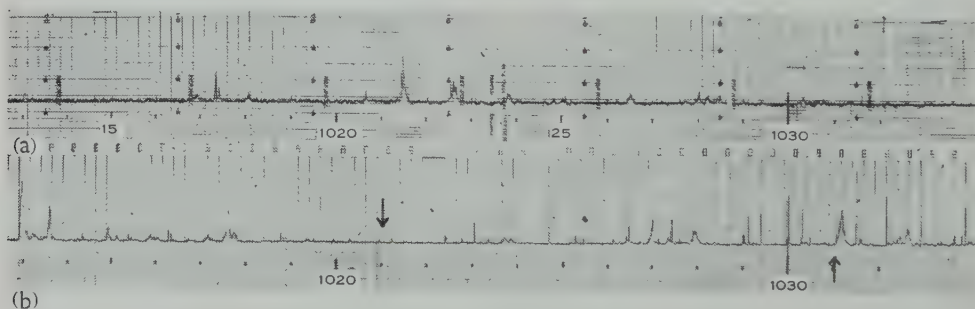


Fig. 6.—Jupiter radiation on two frequencies at Fleurs, February 26, 1956, (a) 27 Mc/s—single aerial record, (b) 19.6 Mc/s—interferometer record, the time of maximum response is shown by an upward-pointing arrow and the time of minimum response by a downward-pointing arrow.

Figure 6 shows simultaneous records at the two frequencies. A few bursts seem to occur on both frequencies, but the majority do not. The agreement in Figure 6 is probably worse than usual, but during any period of activity common to both frequencies, no more than half the bursts appeared on the two frequencies. We do not know to what extent these differences are due to the ionosphere. From our limited spaced-receiver observations it would appear the bursts on two frequencies at one site differ more than bursts on the one frequency but at spaced sites. The peak intensity on 27 Mc/s was generally only about one-third of that

on 19.6 Mc/s, and on those days on which 27 Mc/s radiation was received the duration of the activity was only about one-half of that on 19.6 Mc/s. It is quite clear that there was a large decrease in the overall activity with an increase of operating frequency of only 7.4 Mc/s (that is, about one-third of 19.6 Mc/s).

The results obtained at 14 Mc/s were seriously limited by frequent severe station interference near this frequency. However, unless the association between days of activity on 14 and 19.6 Mc/s is worse than between 19.6 and 27 Mc/s, it is reasonably certain that both the peak intensity and the mean duration of active periods on 14 Mc/s are lower than on 19.6 Mc/s. This is not thought to be due only to the greater ionospheric attenuation on this frequency. When the 14 Mc/s observations were made, at night-time between November 1955 and January 1956, the *F*-region critical frequency was usually below 5 Mc/s near the time of transit of Jupiter, and with an angle of incidence on the ionosphere of about 45° the attenuation of 14 Mc/s radiation should be small. This conclusion is supported by observations of solar bursts made with about the same angle of incidence but during the day-time when critical frequencies were higher. Peak intensities of the *solar bursts* on 14 Mc/s were approximately twice those on 19.6 Mc/s.

TABLE 2
SPECTRAL OBSERVATIONS OF JUPITER RADIATION

Frequency	Results	
27 Mc/s	Peak intensity	$\approx \frac{1}{3}$ of 19.6 Mc/s
	Mean duration of active periods	$< \frac{1}{2}$ of 19.6 Mc/s
	Frequency of occurrence (day by day)	$\approx \frac{1}{2}$ of 19.6 Mc/s
14 Mc/s	Peak intensity	< 19.6 Mc/s
	Mean duration of active periods	< 19.6 Mc/s

The spectral information is summarized in Table 2. The table indicates a maximum in the level of activity between 14 and 27 Mc/s; the maximum is probably below 20 Mc/s.

We have no evidence for progressive changes in the frequency structure of the radiation, e.g. frequency drifts, of the type often found in solar noise.

(e) Polarization

The only good record with the polarization apparatus was obtained on January 24, 1956. The longitude of the central meridian was then near 300° in System II and the most active source was being observed (see below). The downcoming wave was then approximately circularly polarized with a right-handed sense of rotation (adopting the radio-astronomical convention of looking along the direction of travel, in this case from Jupiter to the Earth). The two transverse components, in and at right-angles to the plane of incidence, were equal, within a factor of 3 to 1 in amplitude, while their phase difference was $90 \pm 25^\circ$.

This result agrees with that of Franklin and Burke (1956) from observations at Washington. If the sense of rotation is indeed the same when observed in both the northern and southern hemispheres (and obviously more observations are required), the terrestrial ionosphere cannot be the medium which impresses circular polarization on the Jupiter radiation.

(f) Variation of Activity with the Rotation of Jupiter

In the previous paper (Shain 1956) it was shown that Jupiter radiation in 1951 had a strong tendency to recur at intervals of a rotation period, an indication that the active centres on the planet were of small area and persisted for more than one rotation. This recurrence tendency was still prominent in 1955-56. The earlier part of the combined Sydney and Washington data has been discussed by Alexander (1956). Here we consider only our own observations and all the results are included. Washington results give a similar picture (Franklin and Burke 1956).

The recurrence tendency is shown by the arrangement of the observational data in Figure 7. The times of reception of Jupiter radiation have been converted to central meridian longitudes in System II. In this system 360° corresponds to the rotation period of $9^h 55^m 40^s$.

The tendency for the lines to fall under one another is very marked, and it can be concluded that the principal source of the radio disturbance near longitude 320° was revolving with a speed close to that of System II, and persisted for the full 7 months, or some 500 revolutions.

To estimate the rotation period more closely, the central time of the activity on each day was noted and the dashed line in Figure 7 shows the least squares line of best fit to all the data. It corresponds to a rotation period of $9^h 55^m 34^s$ (6^s faster than System II), with a standard deviation of 9^s . However, it will be shown in Figure 8 (Section III (g) below) that the data can be divided into three groups, and, if the analysis is restricted to the main source, which is observed over only 60° of longitude, the rotation period is found to be $9^h 55^m 30^s$ with a standard deviation of only 3^s . The source was therefore moving significantly faster than System II. The speed of rotation was apparently slower than that of the source observed in 1951 (rotation period $9^h 55^m 13^s$, Shain 1956) but as the standard deviation for a similar analysis of the 1951 data is 30^s (a rough estimate of the probable error was previously given as $\pm 5^s$) the difference may not be significant.

There are suggestions of two disturbances following System I (rotation period $9^h 50^m 30^s$) from September to November, each producing an increase in overall activity as it crosses the line of the principal source, but the data are too few to be conclusive.

The 1951 results indicated a possible connexion between the source of the radio disturbance and a white spot (referred to as DE) at the southern edge of the South Temperate Belt (S.T.B.). Thin lines in Figure 7 show the positions, during the current period, of this and two other similar spots, in the S.T.B., taken from optical observations quoted by Alexander (1956). The arrow at the bottom of the figure shows the position of the centre of the Red Spot, whose

longitude was constant over the period. There is no obvious connexion between possible sources of the radiation and any of the S.T.B. spots, and certainly they cannot be the principal source. The proximity of the Red Spot to the principal source suggests a coincidence, but there are systematic differences in longitude

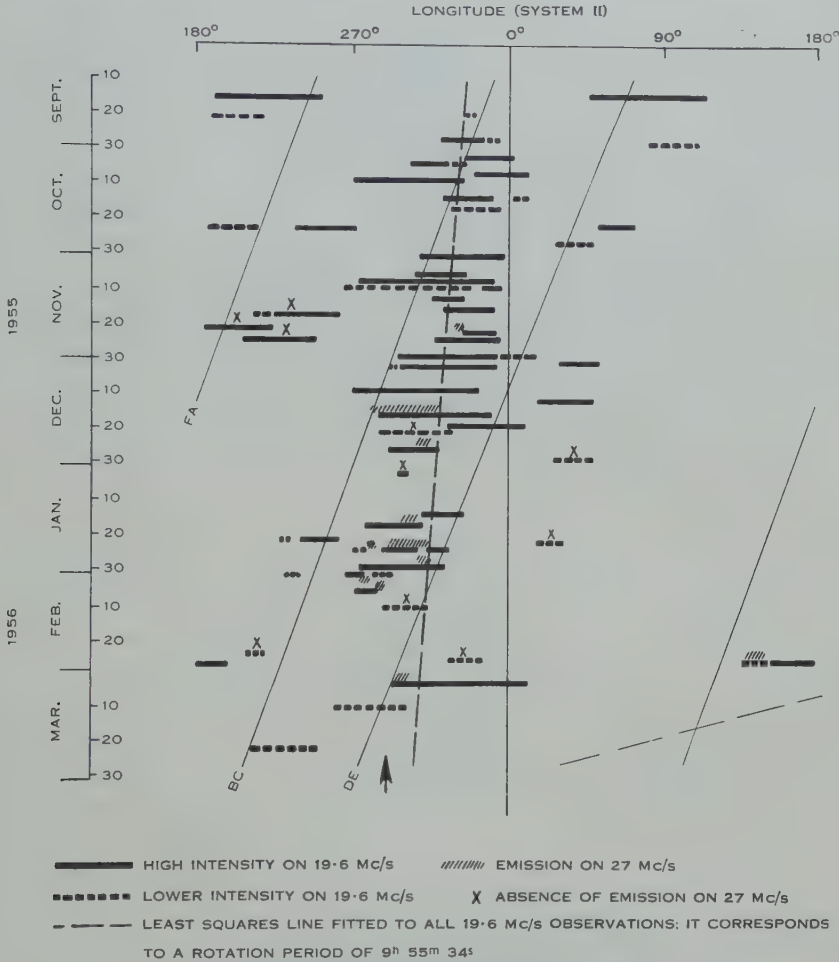


Fig. 7.—Diagram showing central meridian longitude (System II) of Jupiter at times of observed radio emission. The positions of three white spots, FA, BC, and DE, are shown by thin lines—and the Red Spot by an arrow \uparrow . The slope on this diagram of a source moving with System I speed is shown — — — —.

and, more important, in rotation periods and it is most probable that the near coincidence is fortuitous. It is worth noting that the Red Spot was definitely not the main source of the 1950–51 radio noise.

In an attempt to obtain some positive identification of the radio source, plans were made for simultaneous visual observations of Jupiter at times of radio activity, but exceptionally bad weather prevented any useful work along these lines.

(g) Angular Spread of the Emitted Radiation

Following the procedure used previously with the 1951 observations, all the lines for the 19.6 Mc/s observations in Figure 7 were superimposed, with allowance for the slight drift in longitude, to give the histogram of Figure 8.

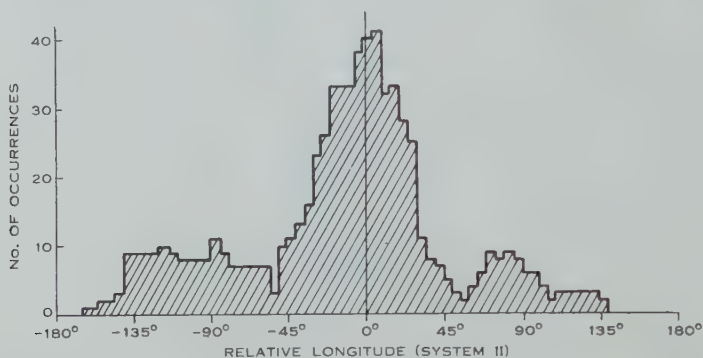


Fig. 8.—Variation of frequency of occurrence of 19.6 Mc/s radiation from Jupiter with central meridian longitude (System II) measured relative to the least squares line of Figure 7.

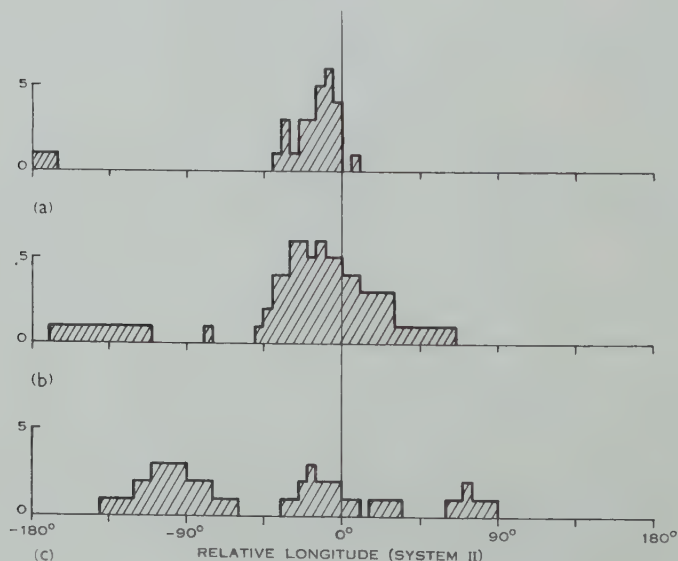


Fig. 9.—Variation of the frequency of occurrence of Jupiter radiation with central meridian longitude measured relative to the least squares line of Figure 7, (a) 27 Mc/s, (b) 19.6 Mc/s on same days as 27 Mc/s radiation was observed, (c) 19.6 Mc/s on days when there was no observable radiation on 27 Mc/s.

This figure shows the frequency of occurrence of Jupiter radiation for 5° intervals of longitude, measured relative to the dashed line in Figure 7. The emission from what we have taken to be the principal source can be seen to fall almost to zero at longitudes 45° from the maximum. The other two peaks, at relative

longitudes -100° and $+80^\circ$, are probably due to independent sources. Figure 8 gives something rather equivalent to the average "polar diagram" of the emission from these sources. On the reasonable assumption that the vertical scale is equivalent to a scale of power, the "beamwidth" of the sources of radiation would be about $\pm 30^\circ$ to half-power.

Histograms were compiled in a similar way to compare the 19.6 and 27 Mc/s results. Figures 9 (a) and 9 (b) show histograms for 27 and 19.6 Mc/s on days on which 27 Mc/s radiation was received, and Figure 9 (c) for 19.6 Mc/s for days when there was no observable radiation on 27 Mc/s. The longitude scale is the same as that in Figure 8. The absence in Figures 9 (a) and 9 (b) of the prominent peaks at -100° and $+80^\circ$ is the result, previously noted above, that some sources radiate on 19.6 but not on 27 Mc/s. This supports the conclusion that the three peaks in Figure 8 are associated with independent sources of radiation.

When attention is restricted to the principal source, common to both frequencies, we see clearly from Figure 9 that the angular spread of its radiation is much narrower on 27 Mc/s than on 19.6 Mc/s. The result is also found that 27 Mc/s activity occurred, on the average, only during the early part of 19.6 Mc/s activity. That is, the histogram in Figure 9 (a) is significantly displaced with respect to the histogram of the principal source in Figure 9 (b).

IV. DISCUSSION

The main facts uncovered in the present investigation may be summarized as follows :

(i) The radiation lasted for periods of about half an hour and was made up of short bursts lasting for times of the order of 1 sec. Over several months Jupiter was active for only about 5 per cent. of the time.

(ii) The intensities of the bursts were very high, attaining a peak of about $10^{-10} \text{ W m}^{-2} (\text{c/s})^{-1}$.

(iii) Marked differences were observed between records of radiation made simultaneously at stations 25 km apart.

(iv) The intensity and frequency of occurrence was about 2 or 3 times greater on 19.6 Mc/s than on 14 or 27 Mc/s.

(v) From one record only in Sydney (with confirmation from Washington) the radiation was approximately circularly polarized.

(vi) During the period of the observations there were three sources of 19.6 Mc/s radiation ; none of these can be identified with visual features. Only one of these sources, the most active, radiated appreciably on 27 Mc/s as well as on 19.6 Mc/s. Its rotation period was $9^{\text{h}} 55^{\text{m}} 30^{\text{s}}$, locating it outside the equatorial regions of the planet.

(vii) The angular spread of the radiation was 60° at 19.6 Mc/s but only about 30° at 27 Mc/s. The longitudes of the central meridian for the 27 Mc/s activity were asymmetrically disposed with respect to those for 19.6 Mc/s.

Although it is clear from the spaced-receiver experiment that the terrestrial ionosphere has a pronounced effect on the short-term characteristics of the

received radiation, we consider that the items listed above (with the exception of (iii)) describe the general variation of the activity of the sources on Jupiter. In the light of these data we may consider possible mechanisms of the origin of the radiation.

The rapid variations in intensity and the extremely high brightness temperatures required—above 10^{13} °K, even on the unlikely assumption that the whole disk emits uniformly—make thermal radiation most unlikely. The observed Jupiter radiation differs from terrestrial atmospherics and cosmic noise particularly in its very restricted spectrum. There has been no sign of Jupiter radiation at 38 Mc/s (Burke and Franklin 1955; Smith 1955), a frequency only twice that at which the intensity is a maximum. Such a high frequency cut-off is much more rapid than that found with either of these sources of noise.

Radiation from the Sun, however, sometimes resembles Jupiter radiation in its great variability and spectral concentration, and similar origins are possible. This would require on Jupiter an ionized region with a plasma frequency, or possibly a gyrofrequency, of about 20 Mc/s. The polarization results imply a magnetic field. The ionized region could be excited into oscillation by some form of electrical discharge, or the passage of a shock wave (perhaps originating in some volcanic disturbance).

The restricted spectrum and the constancy of the frequency of maximum intensity over long periods of time suggest a complete layer of ionization surrounding Jupiter, similar to the terrestrial ionosphere. Solar radiation is the most likely source of such ionization, although, at first glance, it would be expected that critical frequencies would be considerably lower than on the Earth. If plasma oscillations are responsible for the radiation, the decay times should approximate ν^{-1} , where ν is the collision frequency (Jaeger and Westfold 1949). The observed decay times of about 0.5 sec suggest that the ionized layer is relatively higher on Jupiter than on the Earth ($\nu^{-1}=10^{-3}$ in the F region).

It is difficult to explain the restricted angular spread of the emitted radiation and its further decrease with increasing frequency. While effects from refraction and absorption by ionized material above the level of emission could produce a narrower beam, they would simultaneously cause an opposite frequency variation, a beam widening towards higher frequencies. This effect might be offset if the higher frequencies originated at greater depths in Jupiter's atmosphere. It should also be noted that, if the angular spread of the radio emission is the same in latitude as in longitude, it is possible that the two sources which did not appear to have radiated on 27 Mc/s are in higher latitudes than the principal source, the narrower "beamwidth" on 27 Mc/s cutting off emission in the direction of the Earth.

Related to the question of the narrower beamwidth at 27 Mc/s is the asymmetry of the 27 Mc/s radiation pattern with respect to 19.6 Mc/s. One important possibility which must be considered is that the 27 Mc/s observations give the true longitude of the source, and the broader, asymmetrical beam at 19.6 Mc/s is due to differing propagation conditions before and after noon on Jupiter. This would affect any optical identifications, but the matter can only be resolved by extensive observations at more than two frequencies.

V. CONCLUSION

The great gap in our present knowledge of Jupiter radio emission is still the lack of any definite identification of the source of the noise with visual features, so that there is no direct tie-up between radio and visual observations. Simultaneous visual and radio observations are the outstanding requirement. It is short-term (order of half an hour) changes in the visual appearance which must be looked for. It must be remembered that the radio source may be situated high in Jupiter's atmosphere so that direct visual observation of the source may prove impossible.

As regards radio observations, the important points requiring elucidation are the magnitude of the effect of the terrestrial ionosphere and the nature of the dynamic spectra of the bursts. On a longer-term basis, observations over an appreciable fraction of a sunspot cycle may help to determine the importance of solar photo-ionization in Jupiter's atmosphere, but such observations must be calibrated for intensity more accurately than has been done in the past.

VI. ACKNOWLEDGMENTS

The authors are grateful to Mr. L. F. Clague for his help in the construction and operation of the equipment, and to Dr. J. L. Pawsey for helpful discussions. Although bad weather prevented the simultaneous radio and visual observations of Jupiter, thanks are due to Dr. A. J. Way and Mr. Harley Wood for their cooperation in planning such work.

VII. REFERENCES

- ALEXANDER, A. F. O'D. (1956).—*J. Brit. Astr. Ass.* **66** : 208.
BURKE, B. F., and FRANKLIN, K. L. (1955).—*J. Geophys. Res.* **60** : 213.
FRANKLIN, K. L., and BURKE, B. F. (1956).—*Astr. J.* **61** : 177.
HEWISH, A. (1951).—*Proc. Roy. Soc. A* **209** : 81.
JAEGER, J. C., and WESTFOLD, K. C. (1949).—*Aust. J. Sci. Res. A* **2** : 322.
KRAUS, J. D. (1956).—*Sky & Telesc.* **15** : 358.
SHAIN, C. A. (1955).—*Nature* **176** : 836.
SHAIN, C. A. (1956).—*Aust. J. Phys.* **9** : 61.
SMITH, F. G. (1955).—*Observatory* **75** : 252.

GAIN MEASUREMENTS OF LARGE AERIALS USED IN INTERFEROMETER AND CROSS-TYPE RADIO TELESCOPES

By A. G. LITTLE*

[Manuscript received September 20, 1957]

Summary

A method has been developed for measuring the gain of large interferometer and cross-type radio telescope aerials. Use is made of the strong discrete radio sources, whose intensity need not be known, to allow comparison of the gains of the aerials with that of a standard.

The aerials of the 3.5 m Mills Cross radio telescope at Sydney have been calibrated in this way, using a dipole with a plane reflector as a standard. The five radio sources, Pictor-A (IAU 05S4A), Hydra-A (IAU 09S1A), Hercules-A (IAU 16N0A), Virgo-A (IAU 12N1A), Taurus-A (IAU 05N2A), were used in the calibration and the flux densities of these sources were then derived.

I. INTRODUCTION

Although many improvements in radio telescopes have been made over the last few years, the accurate measurement of the flux densities of discrete sources still presents a challenging technical problem. A solution to this can only be achieved by careful attention to the calibration of both receiving and aerial systems. The calibration of the latter is particularly difficult.

If the aerial is large, the resulting narrow beam and high gain combine to give a high signal-to-noise ratio and good discrimination against sources other than the one under observation. For accurate flux density measurements both these factors are important, but for large aerials they are counterbalanced by the difficulty of determining the gain with precision. Computations of the gain of such aerials are often unsatisfactory because of difficulties in accounting for the effects of ground reflections, losses, and mutual impedances, whilst direct gain measurement by conventional methods is a formidable task and is quite often impossible. If, on the other hand, small aerials are used, the situation is reversed. The gain may be calculated and measured quite accurately but the signal-to-noise ratio is poor and the low angular resolution results in confusion between sources, so that flux density measurements are again unreliable.

However, as will be shown, it is possible to calibrate large aerials in terms of a small one whilst still retaining the best features of each system. The method is particularly applicable to the calibration of those interferometer or cross-type aerial systems which have two or more large aerials. These two and the smaller standard aerial can be used to form three interferometer pairs, and, from three separate observations of a discrete source, the gains of the two large aerials

* Division of Radiophysics, C.S.I.R.O., University Grounds, Chippendale, N.S.W.

can then be compared with the known gain of the smaller. Since it is a comparison method, neither the absolute strength of the source nor the absolute calibration of the receiver need be known. Thus, used in conjunction with accurate receiver calibration techniques, this method allows accurate flux density measurements to be made.

The aerials of the 3.5 m Mills Cross (Mills *et al.* 1958, in press) at Sydney have been calibrated in this way using a half-wave dipole above a plane reflector as a standard. Five of the strongest radio sources were used in the calibration, and from the measured values of the gain the flux densities of these sources have been calculated.

II. METHOD

Consider the output of a phase-switched two-aerial interferometer with a square-law detector. The increment of receiver output power due to a source of flux density S in each aerial beam is given by

$$\Delta P_1 = KS \sqrt{\{G_A(\theta, \Phi) G_B(\theta, \Phi)\}}, \dots\dots\dots (1)$$

where K is a constant and the square-root factor is the geometric mean of the aerial power gains. θ and Φ are the coordinates of the source relative to the aerials, measured in the horizontal and vertical planes respectively.

If a third aerial of known gain $G_C(\theta, \Phi)$ is now paired with each of the aerials, two more equations are obtained,

$$\Delta P_2 = KS \sqrt{\{G_A(\theta, \Phi) G_C(\theta, \Phi)\}}, \dots\dots\dots (2)$$

$$\Delta P_3 = KS \sqrt{\{G_B(\theta, \Phi) G_C(\theta, \Phi)\}}. \dots\dots\dots (3)$$

The measurable quantities ΔP_1 , ΔP_2 , ΔP_3 may be expressed in terms of equivalent diode noise generator currents I_1 , I_2 , I_3 , and using (2) and (3) with (1) we have

$$G_A(\theta, \Phi) = G_C(\theta, \Phi) (I_1/I_3)^2, \dots\dots\dots (4)$$

$$G_B(\theta, \Phi) = G_C(\theta, \Phi) (I_1/I_2)^2. \dots\dots\dots (5)$$

Thus the gains of two of the aerials are obtained in terms of the measurable power ratios (I_1/I_3) , (I_1/I_2) , and the gain of the third aerial.

The method may be applied to both steerable and transit type instruments, but when applied to the latter we need consider only one of the coordinates, the elevation Φ , as variable; the coordinate θ is fixed usually at $\theta=0$, where the gain in this direction is a maximum. It is essential to measure the gain as a function of Φ , the elevation, because it is not necessarily constant. Tilttable aerials, for example, are affected by ground reflections for low elevations, whilst fixed arrays of dipoles are subject to gain changes as the phasing of the elements is altered to swing the beam from the vertical. Thus, when this method of aerial calibration is applied to transit instruments, a calibration curve as a function of Φ can be obtained by using several sources each at a different elevation.

As already noted, the equations given above are independent of the strength of the source and the absolute calibration of both the receiver and noise generator.

Nevertheless, it is desirable to use strong sources for good signal-to-noise ratios. At the same time, the sources should be much smaller in angular size than the narrowest aerial beam, otherwise uncertain corrections have to be applied. The detector must be a square-law one for the above equations to hold; if a linear detector is used, the equations have to be modified to include the effects of the background radiation on the amplitude of the recorded signal. This correction would be different for different combinations of aerials.

In the foregoing, it has been assumed that only one source at a time is present in the aerial beam, that is, there is no confusion. Although this is not true of the smaller standard aerial, enough discrimination can be provided by the other aerial of the interferometer pair if it is sufficiently directive. Consequently, the use of the method is restricted to the calibration of pairs of large aerials such as are found in the large cross or interferometer-type radio telescopes. At short wavelengths, where the Sun is by far the brightest object in the sky, smaller aerials could be calibrated by the method, using the Sun without danger of confusion effects, although for these aerials conventional methods of calibration may be more suitable.

The accuracy of this method is limited by the signal-to-noise ratio. An accuracy of 5 per cent. in the noise generator current ratios is required for the gain to be obtained to 10 per cent., and in the presence of noise such accuracies can be difficult to achieve. High signal-to-noise ratios are therefore required. A further difficulty is introduced by ionospheric scintillations, which effectively add to the noise fluctuations and in serious cases can even render an observation useless.

These difficulties can be overcome by taking a sufficient number of observations.

III. APPLICATION TO A CROSS AERIAL

The method has been applied to the calibration of the 3.5 m Mills Cross aerial at Sydney. The experimental arrangement is shown in Figure 1. A half-wave dipole and plane reflector were placed at a distance from the centre of the Cross and pairs of aerials were connected to a phase-switched receiver to be described elsewhere (Mills *et al.* 1958, in press). The output of this was displayed on a Speedomax recording millivoltmeter.

The choice of a standard aerial appears to be limited to either a dipole-reflector combination or a horn. Seeger (1956) has shown what can be achieved with a horn at short wavelengths, but when used at a long wavelength they are clumsy and are susceptible to ground reflection effects, which in the case of the dipole with reflector can be eliminated altogether. Hence this type of aerial has been used here. The dipole was made of $\frac{1}{4}$ in. diameter copper tubing, spaced 0.21 wavelength above a wire mesh reflector which measured approximately 2 by $1\frac{1}{2}$ wavelengths. This whole structure was placed on the ground, which then acts as an infinite extension of the reflecting screen. From the conventional definition for directivity (see for example Schelkunoff and Friis

(1952)) the following expression for the gain of this system in the direction Φ can be derived

$$g_A = 4 \frac{R_D}{R_A} g_D \sin^2 \left(\frac{2\pi a}{\lambda} \cos \Phi \right), \dots\dots\dots (6)$$

where the g 's are the isotropic gains and the R 's the aerial input resistances. The subscript D refers to a dipole in free space and A to a dipole at a distance " a " wavelengths above a plane reflector. λ is the wavelength. The angle Φ is measured from the vertical in the north-south plane which is normal to the axis of the dipole. The response of the dipole in this direction is a constant. The calculated maximum value for g_A was 6.0, neglecting ohmic losses, which amount to less than $\frac{1}{2}$ per cent. for the aerial used.

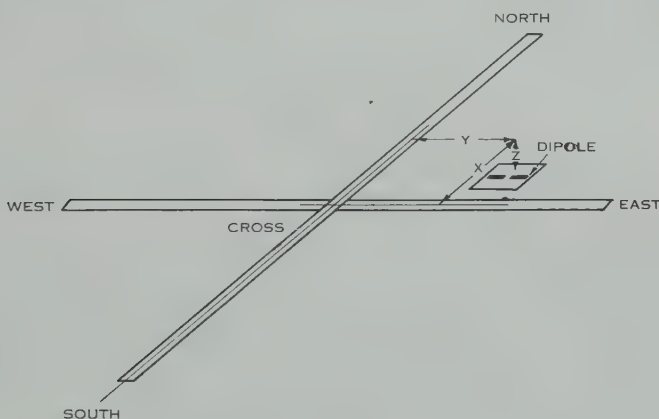


Fig. 1.—Experimental arrangement of aerials.

The loss in the cable connecting the dipole to the receiver was found to be 2.3 ± 0.1 dB by an insertion loss measurement made between a matched load and a matched diode noise generator.

Two dipoles were constructed. One was half a wavelength long and was matched by a Pi network of known loss at a point half a wavelength from the input terminals. The other was matched by shortening the dipole and by placing balanced series reactances at the dipole terminals. This dipole was thus shorter than a half wavelength by 6 per cent.

The relative gains of these two were checked by observations of the quiet Sun using one dipole at a time with the north-south arm of the Cross. An interference pattern was obtained in the way to be described later and the dipoles were interchanged at transit above the same reflector.

The gains were found to be the same to within 3 per cent., and for mechanical reasons the shortened dipole was used in the following observations.

The three aerial combinations possible with this arrangement produce three distinct types of record as shown in Figure 2. Because of the spacing between the dipole and the centre of the Cross, an interference pattern is produced. This is shown in case (a), where a large number of fringes are observed due to the

broad east-west pattern of the north-south aerial. In case (b) the narrow pattern of the east-west aerial allows only one fringe to be observed. Case (c) is the normal pencil-beam record of the Cross.

For these measurements it was decided to use the strong radio sources, Pictor-A (IAU 05S4A), Hydra-A (IAU 09S1A), Hercules-A (IAU 16N0A), Virgo-A (IAU 12N1A), and Taurus-A (IAU 05N2A), since these culminate at different zenith angles and therefore allow the gain to be obtained as a function of elevation. The use of the Sun is excluded because it is larger in angular size than the aerial beams used in the Cross.

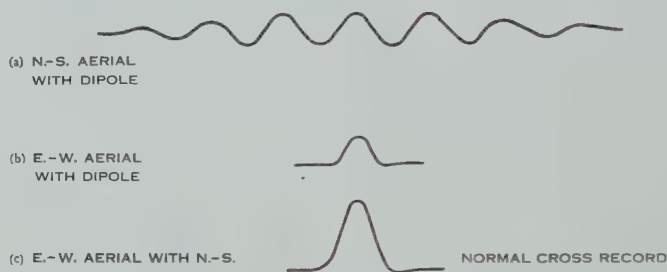


Fig. 2.—East-west patterns produced by different aerial combinations.

Three separate records are then required for each of these sources, but because the cross is a transit instrument it was impossible to get them all on one day. However, it was possible to get two records without losing much information, by first setting up the north-south aerial with the dipole and then near the time of transit changing over to one or other of the remaining aerial combinations for 10 min.

TABLE 1

TOTAL NUMBERS OF RECORDS TAKEN FOR EACH AERIAL COMBINATION AND SOURCE

Source	Aerial Combination		
	Cross	E.-W. Dipole	N.-S. Dipole
Pictor-A, 05S4A	3	3	4
Hydra-A, 09S1A	3	6	4
Hercules-A, 16N0A	6	5	2
Virgo-A, 12N1A	5	4	7
Taurus-A, 05N2A	3	6	3

The numbers of records taken for each aerial combination and source are given in Table 1.

A recording of the source Taurus-A is reproduced in Figure 3. For this case both the north-south with dipole and Cross aerial combinations were used. Noise fluctuations are not troublesome for this source.

A second sample record is shown in Figure 4, and was taken of the source Hydra-A using the north-south with dipole and east-west with dipole aerial combinations. Noise fluctuations are in this case more important. This is particularly so for the single peak observed using the east-west aerial with the dipole. Also, this part of the record had barely enough time to establish a base level from which to measure the height of the peak. As a result of these two factors additional records of the type shown in Figure 5 were taken for all sources.

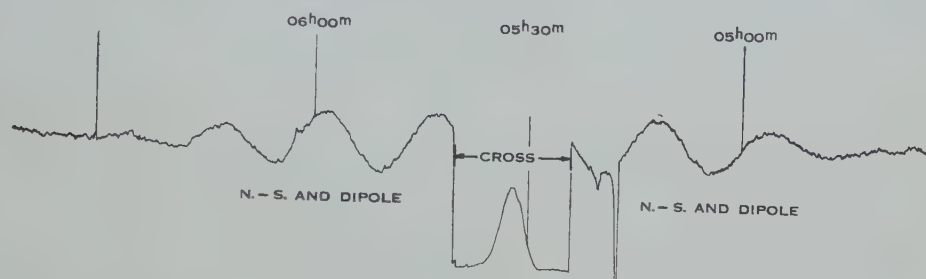


Fig. 3.—Records taken on August 21, 1956 of the discrete radio source Taurus-A using the aerial combinations as shown. The large negative spike at about 0520 is due to interference.

Those illustrated were taken on consecutive days of the Hydra-A source using only the east-west with dipole aerial combination. Some of the variations are due to scintillations and some to noise. Nevertheless, the smoothed amplitudes of these records do not differ by more than 10 per cent.

In the absence of scintillations, the normal Cross records, such as shown in Figure 3, are reproducible from day to day with an accuracy of 3–4 per cent.

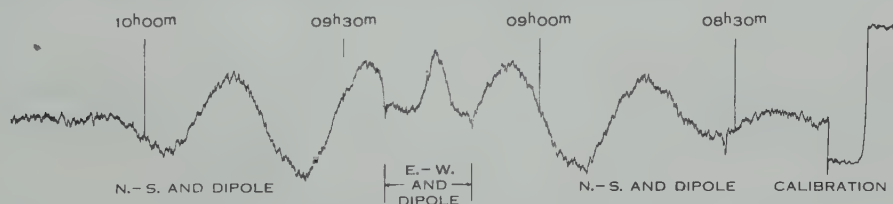


Fig. 4.—Records taken on May 15, 1956 of the discrete source Hydra-A using the aerial combinations as shown.

From all these records it can be seen that the signal-to-noise ratios obtained during these measurements are adequate for calibration purposes, without having to make large numbers of observations. One of the sources, Hercules-A, showed confusion effects. These became apparent when different readings were obtained for different positions of the test dipole with respect to the Cross arms. This source is in a region of strong galactic radiation, and the effect was due to confusion between local concentrations in this radiation and the source. However, such confusion was reduced to negligible proportions when the position of the test dipole was fixed in the position $Y=15.8\lambda$, $X=11\lambda$, $Z=-0.296\lambda$. The coordinate system is shown in Figure 1.

In general, confusion effects due to extended sources may be reduced by using sufficiently large aerial spacings.

All the measurements were taken with the dipole in the position just given, with the exception of those of Pictor-A, which had been completed with the dipole at $Y=13.4\lambda$, $X=1.85\lambda$, $Z=-0.28\lambda$ before the difficulty with the Hercules-A source arose.

A calibration signal similar to that shown in Figure 4 from a diode noise generator was placed on each record. A square-law detector was used and hence the recorded source amplitudes could then be measured off directly as an equi-

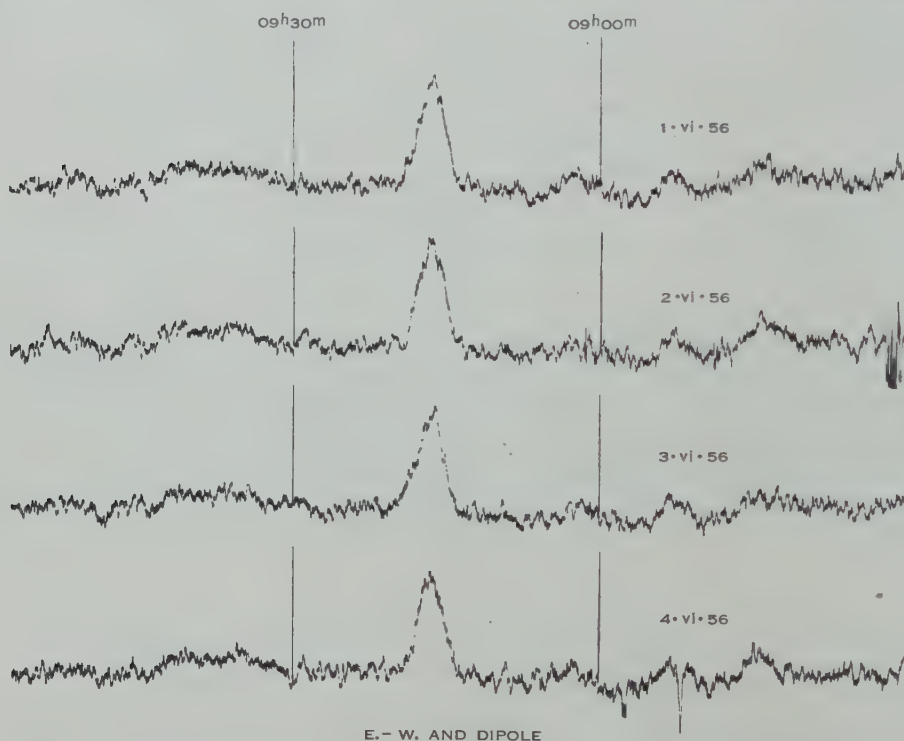


Fig. 5.—Four consecutive records of the source Hydra-A using the east-west aerial with a dipole.

valent noise generator current. A correction was required, however, for the change in amplitude of the interference fringes away from transit. This correction was obtained from a mean of several curves of the observed envelope of the fringe pattern.

From these measurements on the five sources the gains of each arm of the Cross aerial were calculated using equations (4) and (5). Now the gain of the Cross aerial as a whole is given by

$$G_{\text{cross}} = 2\sqrt{G_A G_B}, \dots\dots\dots (7)$$

where G_A and G_B refer to the gains of the north-south and east-west arms respectively.

Substituting from equations (4) and (5)

$$G_{\text{cross}} = 2G_c \frac{I_1^2}{I_2 I_3} \dots\dots\dots (8)$$

The values so obtained are given in Table 2. The probable errors of these values are also given and have been obtained from the observed probable errors of the variables in (8).

These results show that the gain of a cross aerial can be determined at suitable points with a probable error due to random effects of about 5 per cent. Systematic errors are believed to be no greater than 10 per cent.

TABLE 2

MEASURED GAIN OF 3.5 M CROSS AERIAL AT DIFFERENT ELEVATIONS AND FLUX DENSITIES OF SOURCES USED FOR THE CALIBRATION

Source	Elevation	Effective Gain of Cross Aerial over Isotropic	Flux Density ($10^{-24} \text{ W m}^{-2} (\text{c/s})^{-1}$)
Pictor-A, 05S4A	South 78°	820 ± 34	5.7
Hydra-A, 09S1A.. ..	North 68°	640 ± 28	6.7
Hercules-A, 16N0A	North 51°	390 ± 22	8.9
Virgo-A, 12N1A	North 43½°	297 ± 8	24.3
Taurus-A, 05N2A	North 34°	180 ± 10	23

IV. SOURCE FLUX DENSITIES

Having thus determined the gain of the aerial we are now in a position to calculate the flux densities of the sources used for calibration. To do this, the calibration of the receiver was checked, and the values of the flux densities are given in Table 2.* The values refer to both planes of polarization and have a probable error of 10 per cent.

It appears that these values are higher than those previously quoted, as reference to Pawsey (1955) and Shakeshaft *et al.* (1955) will show. The reason for this could be that early measurements were affected by confusion effects due to the poor resolving power of the instruments. Furthermore, most other instruments have not been calibrated as a function of elevation and hence ground reflection effects could cause some of the differences which are observed. It is felt therefore that these new values are not subject to the same uncertainties.

V. ACKNOWLEDGMENTS

The author wishes to acknowledge the valuable discussions with Mr. B. Y. Mills during the project and the helpful advice and criticism of Dr. J. L. Pawsey in the preparation of the manuscript.

* Baldwin and Smith (1956) have suggested that the Virgo-A source is surrounded by a corona, but our measurements point rather to an extension of the source in the general direction of the jet of the associated nebula M87; this does not have any appreciable effect on the flux density given here.

VI. REFERENCES

- BALDWIN, J. E., and SMITH, F. G. (1956).—*Observatory* **76**: 141-4.
- MILLS, B. Y., LITTLE, A. G., SHERIDAN, K. V., and SLEE, O. B. (1958).—A high resolution radio telescope for use at a wavelength of 3·5 metres. *Proc. Inst. Radio Engrs., N.Y.* (In press.)
- PAWSEY, J. L. (1955).—*Astrophys. J.* **121**: 1-5.
- SCHELKUNOFF, S. A., and FRIIS, H. T. (1952).—"Antenna Theory and Practice." (Wiley: New York.)
- SEEGER, C. L. (1956).—*Bull. Astr. Insts. Netherlands*. **13** (472): 100-4.
- SHAKESHAF, J. R., *et al.* (1955).—*Mem. R. Astr. Soc.* **67**: 97-154.

ANOMALIES IN IONOSONDE RECORDS DUE TO TRAVELLING IONOSPHERIC DISTURBANCES

By L. H. HEISLER*

[*Manuscript received October 3, 1957*]

Summary

Anomalies which frequently appear on ionosonde records of the F region during the passage of travelling disturbances are classified into four main types; and the diurnal and seasonal distribution of their occurrence is discussed.

It is suggested that the type of anomaly appearing on records depends on the ion density distribution at a height of about 200 km, which appears to be an upper bounding surface for the mode of travel of disturbances.

A particular study has been made of winter disturbances. These are found to be so frequent that they affect all ionosonde records obtained during this season. They travel distances of at least 3000 km with fronts possibly broader than 1000 km. Attempted correlation with geomagnetic storminess was unsuccessful.

Information is also presented on similar disturbances observed in North America.

I. INTRODUCTION

The existence of disturbances travelling horizontally in the ionosphere has been established by various workers. Munro (1950, 1953) has studied them extensively on a fixed frequency of 5·8 Mc/s, using a three-station triangulation to determine speed and direction. In order to observe disturbances of ion density at different heights in the ionosphere a programme of $h'f$ recording was initiated at these laboratories during 1951 using a fast sweep panoramic type recorder previously described (Heisler 1955).

This recorder has been in daily operation at Camden (150° 40' E., 34° 03' S.) near Sydney from 0900 to 1400 hr local standard time scanning a frequency range of 1–15 Mc/s in 15 sec, the records being made at 1 min intervals. Their study soon revealed the existence of anomalies which, because of their transient nature, are usually unobserved on records taken at intervals of 10 min or more. They appear most obviously at frequencies near the F_1 and F_2 penetration frequencies. Simultaneous observations on a three-station system for detecting ionospheric movements (Munro 1950) have established that these transient irregularities are always manifestations of travelling ionospheric disturbances (T.I.D.'s).

$h'f$ disturbances in general are much greater in amplitude than those studied by the more sensitive fixed frequency techniques and are often so large that penetration occurs on the corresponding fixed frequency record, thus making

* Radio Research Laboratories (Sydney Laboratory), C.S.I.R.O., c/o Electrical Engineering Department, University of Sydney.

determination of speeds and directions impossible. On the other hand, disturbances observed on fixed frequency records are often too small to be resolved satisfactorily on ionosonde records. However, there are many disturbances which show up simultaneously on both types of record; these cases leave little doubt that the disturbances normally observed on these records differ only in magnitude.

This paper presents information on the diurnal and seasonal distribution of occurrence of ionosonde anomalies observed at Sydney during the years 1952 to 1955. (Corresponding statistical data for the smaller disturbances observed

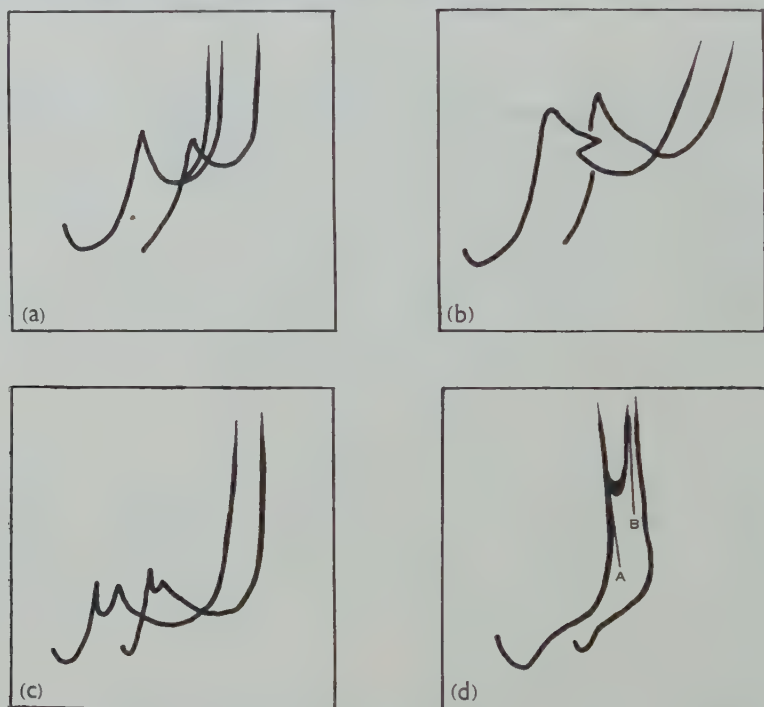


Fig. 1.—Types of anomalies on ionosonde records. (a) Type A anomaly, Camden, 1323 hr, October 25, 1954; (b) type B anomaly, Camden, 1032 hr, November 24, 1951; (c) type C anomaly, Camden, 1114 hr, September 25, 1953; (d) type D anomaly, Camden, 1001 hr, July 31, 1952.

on fixed frequency records are given in a paper by Munro (1958).) In some well-defined cases the amplitude too has been estimated. These show that the change in ion density during the passage of a disturbance is considerable. Analysis of records from all ionosonde stations in Australia indicated that the disturbances travel distances of at least 3000 km with fronts possibly broader than 1000 km.

Data obtained by G. H. Munro during a visit to America in June 1955 are also presented. These show that similar disturbances exist in the northern hemisphere travelling with directions and speeds different from the Australian counterparts.

II. TYPES OF ANOMALIES

Ionosonde records taken at Camden during the years 1952–1955 have been examined and the anomalies observed have been grouped according to four main types. Figure 1 shows tracings of actual records in which these types are immediately evident, and they may be classified as follows:

- (a) Type *A* or “split” type anomaly is marked by a distinct forking of the trace at F_2 penetration frequency. This is very similar to a “penetration frequency multiplet” in spread- F phenomena described by Singleton (1957).
- (b) Type *B* or “Z” type anomaly has a fold in the F_2 trace forming a shape like the letter “Z”.
- (c) Type *C* or “double peak” anomaly occurs in the F_1 trace as a cusp-shaped irregularity forming two distinct F_1 peaks.
- (d) Type *D* or “cusp” anomaly appears as a cusp-shaped trace near F_2 penetration frequency.

Type *A* and *B* irregularities have already been explained by Munro (1953) as being due to complex reflections from curved isoionic surfaces in the ionosphere during the passage of a travelling disturbance. In these cases, only a reduction of ion density was evident, the ion density contours producing the complexities being curved concave upwards.

Types *C* and *D* are the cusp anomalies described by Munro and Heisler (1956). Type *C* occurs as a complexity in the F_1 layer, while type *D* occurs in the F_2 layer. Here a disturbance introduces an initial ion density reduction closely followed by an increase in ionization which produces closed contours or “islands” of higher ionization density. The cusp-shaped irregularities are due to non-vertical reflection of ionosonde transmissions from these islands.

III. SEASONAL DISTRIBUTION OF ANOMALIES

In order to obtain an accurate representation of the seasonal distribution of anomalies, the total number of each type occurring during each month of the years 1952–1955 was divided by the total number of recording hours in that month to give a mean hourly rate of occurrence. Figure 2 shows the seasonal distribution and the total number of recording hours each month for the various types of irregularities. It is apparent that type *B* and *D* anomalies are typical summer and winter phenomena respectively, while type *C* is equinoxial. Type *A* appears to have a random distribution throughout the year.

IV. DISTURBANCE ANOMALIES IN VARIOUS ION DENSITY DISTRIBUTIONS

The average seasonal ion density distributions at Camden on which these irregularities appear differ markedly from one another; these are shown in Figures 3, 4, and 5.

The type *D* anomaly occurs on the winter distribution shown in Figure 3, the lower trace on the diagram being the true height distribution of ion density determined by a method due to Kelso (1952). There is no definite F_1 layer and a true F_1 penetration does not occur. The height of maximum ion density is approximately 200 km.

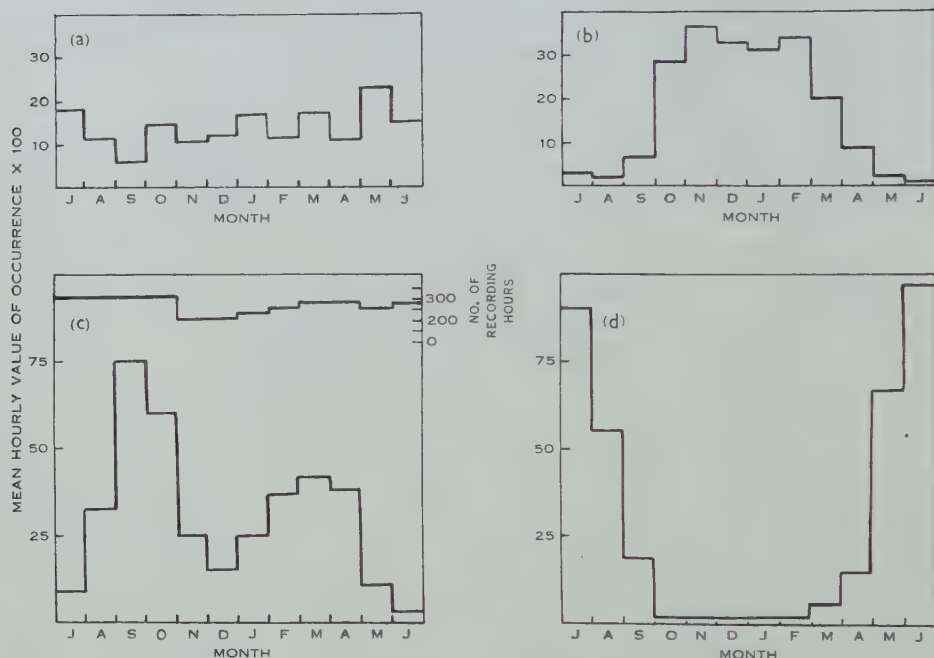


Fig. 2.—Mean hourly rate of occurrence of anomalies each month, July 1952–June 1955.
 (a) Type *A* anomaly ; (b) type *B* anomaly ; (c) type *C* anomaly ; (d) type *D* anomaly.

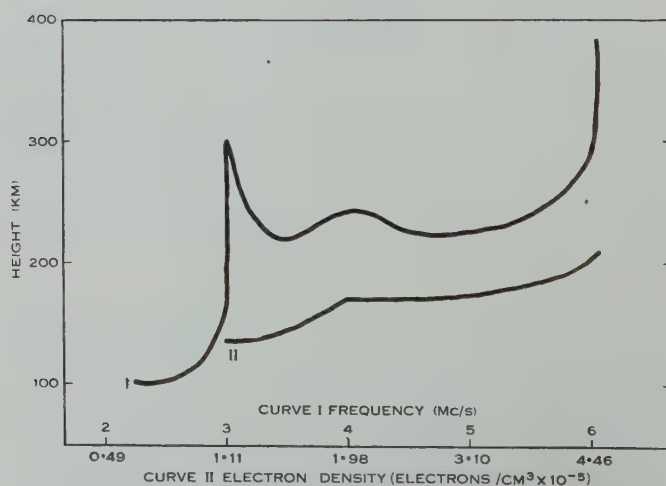


Fig. 3.—Typical ionosonde record (curve I) and true height distribution (curve II) before advent of a type *D* anomaly.
 Camden, 1015 hr, June 25, 1953.

Both virtual and true height equinoxial distributions appear in Figure 4. Here F_1 penetration occurs and a definite maximum value of F_1 electron density exists. Most of the F_2 layer is above 200 km in height with a maximum ion density at approximately 250 km. Type C anomalies appear in the F_1 trace of this distribution but are not seen at greater heights.

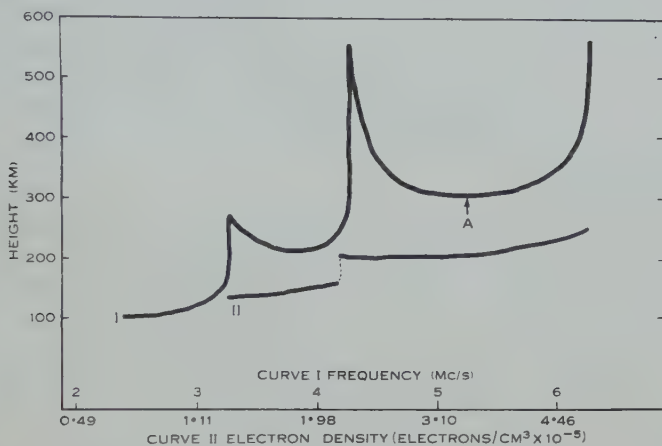


Fig. 4.—Typical ionosonde record (curve I) and true height distribution (curve II) before advent of a type C anomaly. Camden, 1148 hr, September 18, 1954.

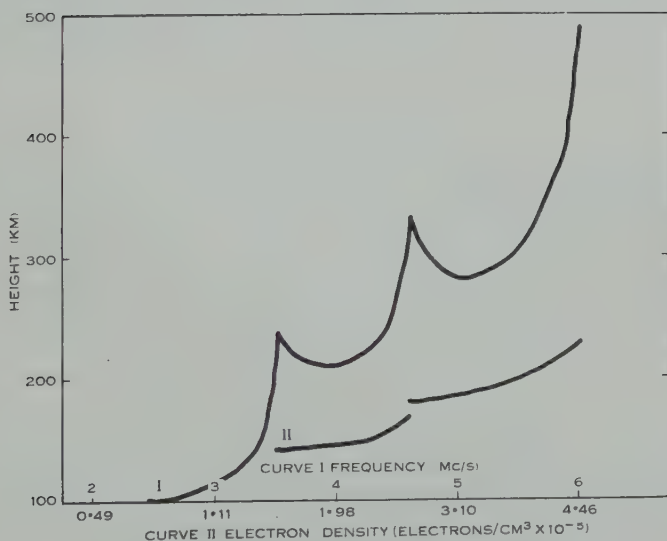


Fig. 5.—Typical ionosonde record (curve I) and true height distribution (curve II) before advent of a type B anomaly. Camden, 1216 hr, October 10, 1954.

Early in October the typical summer distribution develops as represented by Figure 5. This is associated with a sudden change in direction of travel of disturbances as described by Munro (1950). In this case the F_2 layer has a

much greater semi-thickness than formerly, and its maximum again occurs at approximately 250 km. The type *B* anomalies which arise in this distribution rarely occur near the F_2 maximum, and they originate at an approximate height of 200 km on the F_2 trace.

There is a possibility therefore that all disturbances are contained by an upper boundary near 200 km and do not often occur above this height. The normal seasonal change in ionization height-distribution would therefore determine the type of irregularities observed. This would also explain why these are rarely seen on night records when the whole of the F region is well above 200 km. Further research is being undertaken at this laboratory to study this possibility.

Munro (1953) has indicated that complexities on ionosonde records due to travelling disturbance manifestations are more evident in a region where ion density changes slowly with height. This is obvious in Figure 1 where in all cases shown the complexity occurs on a steep portion of the $h'f$ trace.

When a travelling disturbance affects regions where there is a rapid rate of change of ion density with height, as indicated at *A* in Figure 4, there may be no resulting complexity but only a small height change in the trace, probably too small to detect.

Figures 2 (*a*), 2 (*b*), 2 (*c*), and 2 (*d*) also illustrate the high frequency of occurrence of disturbances. For example, disturbances responsible for type *D* anomalies during June occur at the approximate rate of one per hour, and since their duration may be as long as 30 min, records are invariably abnormal during this month.

V. DIURNAL DISTRIBUTION OF ANOMALIES

Figures 6 (*a*), 6 (*b*), 6 (*c*), and 6 (*d*) show the diurnal distribution of anomalies for the years 1952–1955. Type *A* irregularities have a broad maximum during the morning hours, while type *B* have a broad midday maximum. Type *C* anomalies show a marked increase after midday, but unfortunately in this case recording hours have been insufficient to show the time of maximum. Type *D* anomalies show a small broad maximum in the morning hours.

VI. AMPLITUDE OF DISTURBANCES

The type *D* or cusp type anomaly is the most prominent of those discussed above. It is always readily recognizable on records, and during the winter, as already mentioned, it occurs frequently. For this reason it has been chosen as the most suitable type for study at this laboratory.

Figure 1 (*d*) shows this anomaly soon after its first appearance on a series of ionosonde records. The penetration frequency f_A at *A* represents the value N_A to which the ion density has fallen during the initial phase of the disturbance. As the disturbance progresses, the penetration frequency f_B of the cusp as measured at *B* increases, and its maximum value represents the maximum ion density N_B due to the disturbance.

$N_B - N_A$ is a measure of the total ion density change, and, if we assume that the disturbance is sinusoidal in form as suggested by Munro and Heisler (1956), then the percentage amplitude is given by

$$\{(N_B - N_A)/(N_B + N_A)\} \times 100 \text{ or } \{(f_B^2 - f_A^2)/(f_B^2 + f_A^2)\} \times 100.$$

Table 1 shows a number of amplitude values for various disturbances. Some of these amplitudes are quite large, representing as much as 36 per cent. total change in maximum ion density of the F_2 layer.

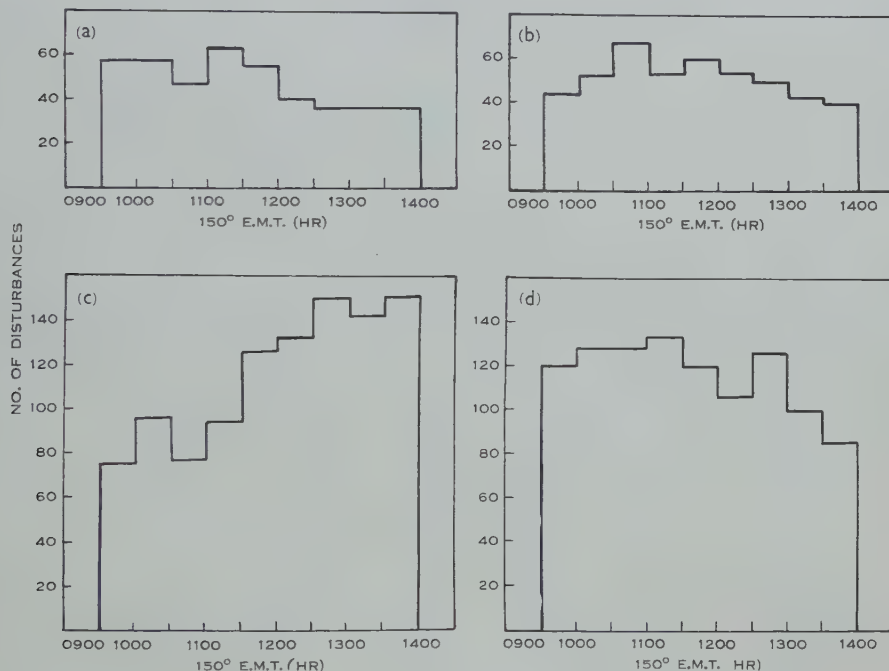


Fig. 6.—Diurnal distribution of anomalies, July 1952–June 1955. (a) Type A; (b) type B; (c) type C; (d) type D.

VII. EXTENT OF DISTURBANCES

Munro (1950) discusses a travelling disturbance of this type which occurred on July 5, 1948, appearing successively on records at Canberra ($35^{\circ} 18' \text{ S.}$, 149° E.), Sydney ($33^{\circ} 52' \text{ S.}$, $151^{\circ} 11' \text{ E.}$), and Brisbane ($27^{\circ} 30' \text{ S.}$, 153° E.). A method of triangulation described in the same paper showed the direction of travel and apparent horizontal speed of this disturbance to be 2° east of north and 10 km/min respectively. This, therefore, was an example of a disturbance travelling horizontally at least 900 km without any definite change in form or velocity.

The large amplitude of type D anomalies discussed above suggested that the extent of these disturbances may be even greater. Consequently, with the kind cooperation of the Ionospheric Prediction Service of the Department of the Interior and of the Bureau of Mineral Resources, Geology and Geophysics

of the Department of National Development, a programme of continuous $h'f$ recordings was arranged for a fortnight during the month of July in 1954, 1955, and 1956. This programme operated at the following stations: Hobart ($42^{\circ} 53' \text{ S.}, 147^{\circ} 51' \text{ E.}$), Canberra ($35^{\circ} 18' \text{ S.}, 149^{\circ} \text{ E.}$), Brisbane ($27^{\circ} 30' \text{ S.}, 153^{\circ} \text{ E.}$), Townsville ($19^{\circ} 10' \text{ S.}, 146^{\circ} 58' \text{ E.}$), and Watheroo ($30^{\circ} 18' \text{ S.}, 115^{\circ} 56.6' \text{ E.}$). By courtesy of the New Zealand Department of Scientific and Industrial Research a few records were also obtained from Christchurch ($43^{\circ} 30' \text{ S.}, 172^{\circ} 30' \text{ E.}$).

TABLE 1
AMPLITUDES OF TYPE *D* DISTURBANCES AT CAMDEN

Date	Time	Minimum f^oF_2 (Mc/s)	Maximum Cusp Penetration Frequency (Mc/s)	Amplitude Ion Density Change (%)
6.vi.53	1044	4.7	5.5	15
10.vi.53	1045	4.9	5.4	10
12.vi.53	1145	5.25	5.75	9
15.vi.53	1123	5.25	5.75	9
16.vi.53	1123	5.1	5.75	12
	1315	5.7	6.5	13
17.vi.53	1241	4.6	5.2	12
19.vi.53	1200	5.0	6.0	18
22.vi.53	1133	5.4	6.0	10
27.vi.53	1228	5.4	6.1	12
28.vi.53	1227	5.4	6.25	14
	1309	5.6	6.25	11
3.vii.53	1108	4.6	5.5	18
	1130	5.3	6.0	12
10.vii.53	0942	4.9	5.25	7
	1132	5.6	6.25	11
16.vii.53	1251	5.0	5.5	9

From a comparison of Canberra, Sydney, and Brisbane records it immediately became obvious that disturbances seen at Canberra appeared at Sydney and Brisbane also, but at different times. Great circle distances and bearings were used to prepare a gnomonic chart of all Australian stations with Canberra as centre; and, assuming that the front of a disturbance was a great circle segment, or a straight line on the chart, it was possible by triangulation of Canberra, Sydney, and Brisbane to obtain directions and speeds of disturbances. Expected times of arrival at Townsville and times of appearance at Hobart were then determined for a large number of disturbances. Examples are shown in Table 2. The time quoted in parentheses is the estimated time of arrival in each case. It can be seen that estimated times and actual times of arrival of a disturbance correspond very closely, and this provides definite evidence of disturbances travelling from Hobart to Townsville, a distance of 3000 km. Moreover, the amplitude of a large disturbance occurring on July 18, 1955 is approximately

18 per cent. at Hobart, Sydney, and Townsville, which implies that for large disturbances at least there is little change in amplitude over the distance travelled.

The stations listed in the table lie approximately in a south-north line, which is the general direction of travel of disturbances in this hemisphere. Distortion of the distance scale on the gnomonic chart in the area considered is very small, and the scale is almost linear in the region bounded by longitudes

TABLE 2

TIME OF ARRIVAL OF DISTURBANCES AT STATIONS AND THEIR ESTIMATED SPEEDS AND DIRECTIONS
All times are 150° E.M.T.

Date	Hobart	Canberra	Sydney	Brisbane	Townsville	Speed (km/min)	Direction (°E. of N.)
14.vii.54 ..	1048 (1048)	1220	1247	1425	*	7.8	7
16.vii.54 ..	1006 (1006)	1152	1215	1346	1600 (1605)	8.4	23
19.vii.54 ..	1152 (1144)	1246	1303	1358	*	11.6	50
21.vii.54 ..	1018 (1024)	1202	1215	1342	1538 (1540)	8.2	350
16.vii.55 ..	† (0834)	1026	1044	1220	1426 (1426)	7.4	356
18.vii.55 ..	1056 (1056)	1242	1258	1424	1534 (1534)	9.0	10
21.vii.55 ..	0922 (0922)	1114	1135	1310	1502 (1502)	7.6	4
	1038 (1038)	1220	1234	1354	1540 (1540)	8.8	353
24.vii.55 ..	1025 (1025)	1124	1135	1234	1339 (1339)	12.4	2
	0917 (0917)	1128	1150	1344	1615 (1615)	5.8	358
25.vii.55 ..	0934 (0934)	1054	1114	1230	1310 (1310)	9.8	30
	1028 (1028)	1134	1155	1314	1342 (1342)	10.2	36

* No record available.

† Ten-minute records only available.

135° E. to 165° E. and latitudes 15° S. to 45° S. Outside this region the scale distortion becomes considerable, and an extension of the triangulation method analysis to include Watheroo and Christchurch is difficult, particularly when there is doubt as to the shape of the disturbance front.

However, the disturbance of July 19, 1954 which passes through Hobart and Brisbane and travels in a direction 50° east of north must have a broad front. On the gnomonic chart this distance is given by the projection of the Hobart-Brisbane line on the disturbance front drawn through Hobart and is approxi-

mately 700 km. This represents the minimum extent of the front, the actual frontal length being probably much greater. A similar disturbance of July 25, 1955, travelling 30° east of north and passing through Hobart and Townsville, has a front of at least 1000 km. In the cases considered disturbance fronts are approximately segments of a great circle.

There are indications that the front of some very large disturbances may be several thousand kilometres and it is intended to investigate this further by means of additional ionosondes placed in an east-west line.

Two disturbances occurring on July 24, 1955 provide an interesting example. These disturbances, travelling in approximately the same direction, almost due north, left Hobart at different times with markedly different speeds. Just before reaching Canberra they crossed over in space but maintained their separate identities to arrive later at Sydney, Brisbane, and Townsville. This means that, two disturbances existing simultaneously in the same medium were travelling at different speeds.

TABLE 3

TIME OF ARRIVAL OF DISTURBANCES AT STATIONS AND THEIR ESTIMATED SPEEDS AND DIRECTIONS
All times are 75° W.M.T.

Date	Ottawa	Morgantown	Derwood	Charlottes- ville	Speed (km/min)	Direction ($^\circ$ E. of N.)
18.i.51	(1012)	1012	1055	1047	5.2	117
	0730	1033	1123	1142	2.7	160
	(0742)					
	(0630)	1130	1215	1241	2.0	176
28.i.51	1030	1313	1321	(1353)	4.5	184
	0717	0842	0913	0921	5.4	160
	(0715)					
	(0616)	1127	1211	1233	2.0	180
	0800	1226	1313	(1325)	2.5	174

During a visit to America in June 1955 G. H. Munro found type *D* anomalies, similar to those discussed above, on northern hemisphere winter ionosonde records. With the kind permission of the Department of Terrestrial Magnetism of the Carnegie Institution of Washington and the Canadian Defence Research Laboratories, records from Morgantown ($37^\circ 7' \text{ N.}, 79^\circ 9' \text{ W.}$), Derwood ($39^\circ 1' \text{ N.}, 77^\circ 2' \text{ W.}$), Charlottesville ($38^\circ 1' \text{ N.}, 78^\circ 5' \text{ W.}$), and Ottawa ($45^\circ 20' \text{ N.}, 75^\circ 43' \text{ W.}$) were analysed, and from three stations speed and direction of travel of disturbances were determined by triangulation methods. Expected time of arrival at the fourth station was then estimated and the results are presented in Table 3. The time quoted in parentheses is the estimated time of arrival of a disturbance. In some cases the estimated time and actual time of arrival correspond very closely. Here again the disturbances are large-scale phenomena, but directions and speeds of travel differ from those observed during the southern hemisphere winter at Sydney. They travel south instead of north, and more slowly.

VIII. CORRELATION WITH MAGNETIC PHENOMENA

Attempts have been made to correlate occurrence of type *D* anomalies with magnetic storm data. Comparisons were made between magnetograms and ionosonde records from Watheroo and no correlation was found between disturbed magnetic records and occurrence of travelling ionospheric disturbances. The extent of these disturbances and the fact that they originate well to the south of Hobart suggested a possible correlation between disturbed days and auroral effects. Here again no correlation was found. Magnetogram records from Macquarie Island ($54^{\circ} 30.2' \text{ S.}$, $158^{\circ} 57' \text{ E.}$) also showed no relationship to disturbed days.

IX. CONCLUSIONS

Ionosonde records taken at Sydney during the years 1952–1955 show transient anomalies that are the result of horizontally travelling disturbances in the ionosphere. These irregularities may be classified into four different types. One of these is random in occurrence, but the other three show definite maxima of occurrence during the winter, summer, and equinoxial periods respectively.

Study of typical true-height ion density distributions for these periods suggests the existence of a possible upper boundary to disturbances in the region of 200 km, and that the position and shape of the distribution with respect to this height determines the type of irregularity on ionosonde records. This study is being continued.

Amplitude of disturbances, in particular of those occurring during winter, is very large, and ion density changes of the order of 30 per cent. are not uncommon.

Disturbances have been shown to travel large distances with wide fronts. Many disturbances travel from Hobart to Townsville, a distance of 3000 km, with no apparent change in amplitude and have fronts of at least 1000 km. It is intended by the use of additional ionosondes to investigate further the shape and extent of disturbance fronts.

Occurrence of disturbances is quite frequent. Equinoxial and winter types occur approximately at the rate of one each hour. Summer disturbances occur at the average rate of one every two and a half hours; but many summer disturbances are probably not seen on records due to E_s blanketing, and it is possible that the frequency of occurrence is much higher than this. Since the duration of disturbances may be as long as 30 min, the ionosphere is almost continuously affected by them. During winter frequent increases in F_2 maximum occur because of disturbances which due to their large extent probably make maximum usable frequencies over any particular circuit higher than the predicted value.

X. ACKNOWLEDGMENTS

This investigation has been carried out in the Electrical Engineering Department of the University of Sydney as part of the research programme sponsored by the Radio Research Board of the Commonwealth Scientific and Industrial Research Organization. The preparation of material by the analysis staff is gratefully acknowledged. The author wishes to express appreciation

to the University of Sydney, in particular to Professor D. M. Myers, for provision of facilities in the Department; and to Dr. D. F. Martyn, F.R.S., the Chief Officer-in-Charge of the Radio Research Laboratories, and to Dr. G. H. Munro, Officer-in-Charge of the Sydney Section, for helpful discussions during the preparation of this paper.

XI. REFERENCES

- HEISLER, L. H. (1955).—*Aust. J. Appl. Sci.* **6**: 1.
MUNRO, G. H. (1950).—*Proc. Roy. Soc. A* **202**: 208.
MUNRO, G. H. (1953).—*Proc. Roy. Soc. A* **219**: 447.
MUNRO, G. H. (1958).—*Aust. J. Phys.* **11**: 91.
MUNRO, G. H., and HEISLER, L. H. (1956).—*Aust. J. Phys.* **9**: 359.
SINGLETON, D. G. (1957).—*Aust. J. Phys.* **10**: 60.

TRAVELLING IONOSPHERIC DISTURBANCES IN THE F REGION

By G. H. MUNRO*

[*Manuscript received October 24, 1957*]

Summary

Observations of the horizontal movements of travelling ionospheric disturbances recorded on a single radio frequency from April 1948 to March 1957 are analysed for seasonal and diurnal variations of occurrence and of direction and speed of travel. Recording was mainly in daylight hours but some limited night results are included. The average number of disturbances recorded was six per day over the period. Observing accuracy and significance of the deduced data are discussed. The frequency of occurrence has a diurnal variation with a marked midday maximum and a seasonal variation with minima at the equinoxes.

The monthly means of direction of travel show a consistent seasonal change from 30° in winter to 120° in summer with a small corresponding change in mean speed from 8 km/min in winter to 7 km/min in summer.

The monthly mean diurnal variation of directions was consistent from 1950 to 1954 but has shown a marked change in the last two summers. Winter directions by day are mainly in the north-east quadrant and have a mean day-time drift towards the north but at night they are predominantly in the north-west quadrant. Summer day-time directions were mainly in the south-east quadrant until December 1956; since then they have tended to the south-west after noon, reverting to the south-east about midnight.

Diurnal variation of speed is of the same order as the seasonal change.

I. INTRODUCTION

Previous papers from this laboratory have described some of the characteristics of travelling disturbances in the F region of the ionosphere (Munro 1950) and some of the effects they produce on the common types of ionospheric records (Munro 1953*b*). Some preliminary information was also given on seasonal and diurnal variations in the direction and rate of travel (Munro 1950, 1953*a*).

Observations have now continued for a period of 9 years and sufficient information is available for significant statistical examination of a number of features of the occurrence and movements of the disturbances.

The present paper deals with horizontal movements of disturbances as observed on a single radio frequency.

II. RECORDING TECHNIQUE

The observing system has been described previously (Munro 1950). Three similar pulsed transmitters operating on a common frequency are spaced about 40 km apart at the apices of a triangle. The transmitters are pulsed successively so that the echo signals from all three may be displayed simultaneously on a

* Radio Research Laboratories (Sydney Laboratory), C.S.I.R.O., c/o Electrical Engineering Department, University of Sydney.

cathode-ray oscilloscope and recorded continuously on photographic film. The effective separation of reflection points is about 20 km, which has proved to be the optimum, since for smaller distances the time difference is too small, whilst for greater distances individual disturbances are not always identifiable at all points. The frequency of 5.80 Mc/s, which has been used most often, has been found very suitable as at least one ray reflected from the F_2 region is normally present. Lower frequencies have occasionally been used, particularly for night observations. Higher frequencies have not been much used, as most of the observations have been in a low sunspot period.

Recording has usually been carried out at two sites. This reduces loss of data due to instrumental failure and also permits emphasis on different features at the two points; for example, at the base station, both E and F echoes are recorded for comparison, while at a field station, where the noise level is lower, the F echoes are recorded in greater detail by the use of higher receiver sensitivity and a more open time-base scale.

Virtual height recording using intensity modulation of the cathode-ray oscilloscope beam has been found the most useful for regular observations. The equipment used is also capable of recording continuously either intensity variations or phase-path changes; these facilities have been used occasionally for special observations not discussed in the present paper.

Recording has been carried out mostly by day from approximately 0700 to 1700 hr. This interval covers the period of most frequent occurrence of disturbances.

Regular observations began in April 1948 and have continued with only minor interruptions.

Night-time observations have been made for limited periods, as will appear in the text.

III. OCCURRENCE OF DISTURBANCES

The types of disturbance manifestations observed on records have been described in previous papers (Munro 1950, 1953*b*). In the statistical examination which follows in this section we shall consider mainly complexities or abrupt changes in virtual height (which are actually very small complexities) as these are interpretable in terms of geometrical optics and are much more precise and reliable indications of disturbances than virtual height maxima and minima or ray "cross-overs", which are sometimes used for detection of movement in the absence of complexities. Variations of intensity may also be used as indications of disturbances, but they are found to be less reliable than complexities because additional variations due to absorption and focusing in lower regions may confuse results. In this section records from only one station are considered and the appearance of a complexity on at least one trace is listed as an occurrence of a disturbance.

(a) *Frequency of Occurrence*

These disturbances are of quite frequent occurrence; for example, the average number recorded per day in the years 1951–1956 was approximately six. The number observed varies considerably from day to day. Ease of

observation depends to some extent on the ionization gradient at the time, but this does not account entirely for the day-to-day fluctuations. There do appear to be "quiet days" when disturbances are either absent or very small, and there may be several such days in succession followed by one or more "active" days. The nature of these variations is shown in Figures 1 (a) and 1 (b) which give the number observed each day during the months of July 1955 (winter) and January 1956 (summer) respectively. The mean number per day in each month for the years 1951–1956 is shown in Figure 1(c).

It will be noticed that there is a rough periodicity present in curves (a) and (b).

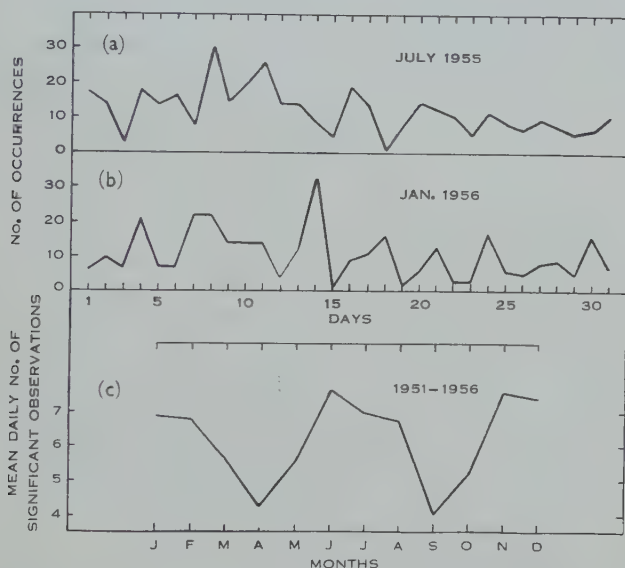


Fig. 1.—Frequency of occurrence of travelling ionospheric disturbances. (a) Day-to-day variations, July 1955; (b) day-to-day variations, Jan. 1956; (c) mean annual variation 1951–1956.

(b) Diurnal Variation of Occurrence

(i) *Day-time.*—The frequency of occurrence of observable disturbances has a marked diurnal variation. The mean variation for each month of the years 1950–1953 is shown in the histograms of Figure 2. The ordinate scale is numbers per hour when conditions were satisfactory for observations. This correction was necessary since the curves would otherwise be distorted because of blanketing by E_s , which has a morning maximum in summer (Harvey 1955). The lower values shown in December and January are the uncorrected ones.

It will be seen that in winter there is a single peak at 1130–1230, but at the equinoxes there are two peaks, before and after midday. This is most marked in September. These peaks are significant since they recur each year, as shown in Figure 3.

Other monthly distributions also repeat from year to year.

(ii) *Night-time*.—Night-time observations have been taken only for a few relatively short periods and the data are insufficient for satisfactory statistical treatment. In general there seems to be an increase in occurrence of disturbances in the middle night hours.

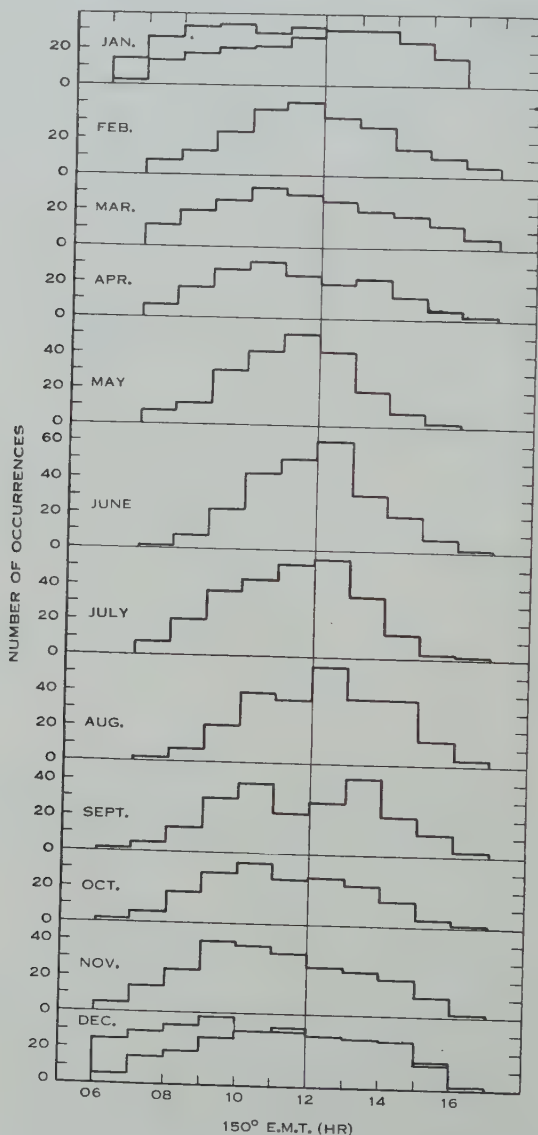


Fig. 2.—Seasonal variation in diurnal distribution of occurrence of disturbances, July 1949 to April 1953.

IV. MOVEMENT OF DISTURBANCES

For the deduction of direction and speed of movement it is necessary that the manifestation of a disturbance on a particular ray ("o" or "x") should be recorded for each of the three transmissions on at least one record. This does

happen in most cases. There must also be no possible ambiguity in correlation of complexities on the different traces. These requirements have been fulfilled for all the data used in this section.

(a) *Treatment of Records and Data*

(i) *Analysis of Records*.—The echo traces corresponding to the three transmitters appear on the film records one above the other so that the time differences in the occurrence of disturbances at the three reflection points are readily observable and unaffected by any errors in actual time of occurrence. Examples of the different types of discontinuity have been shown previously (Munro 1953*b*). To the practised eye they are easily discerned, even when quite small. These discontinuities are far the most frequently observed features, but where none are present peaks or dips in virtual heights and cross-overs of the *o* and *x* ray traces are also used as indicators of disturbances.

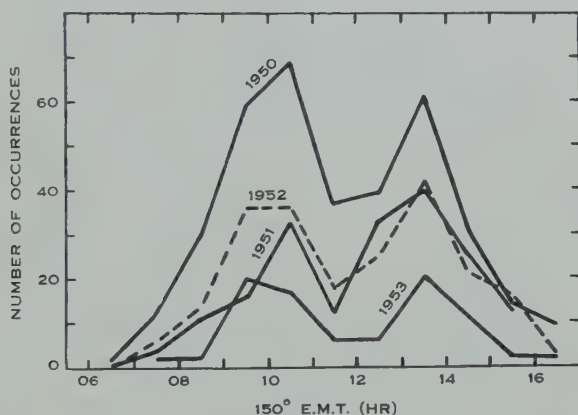


Fig. 3.—Diurnal distribution of occurrence of disturbances in September 1950-1953.

From the time differences in occurrence, the direction and speed of horizontal movement are then determined from tables which have been computed for the particular group of stations, and listed against the earliest time of occurrence.

In the analyses which follow, all such listed data have been used without any form of discrimination.

(ii) *Significance of Data*.—The significance of the analytical results obtained will depend on the quantity and variability of the data, so this aspect requires some detailed consideration.

Assuming that the apparent directions of movement are true directions (this will be discussed more fully in the next section) the accuracy with which they may be determined will depend for a given system on (a) the accuracy with which the time differences of arrival of a disturbance at an observing point may be measured and (b) the speed of travel of the disturbance.

In practice, the times of occurrence of the more clearly defined disturbances are read to the nearest $\frac{1}{4}$ min; others, less clearly defined, to the nearest $\frac{1}{2}$ min

only; whilst there may be a few disturbances where the accuracy is even less. All disturbances are noted, but only the first two classes are used for determining directions and speeds of movements.

With the spacing of stations used at Sydney, there is little chance of wrong correlation of disturbances, as they seldom occur at less than 10-min intervals and the maximum time difference of appearance of a disturbance at two stations is seldom greater than 5 min.

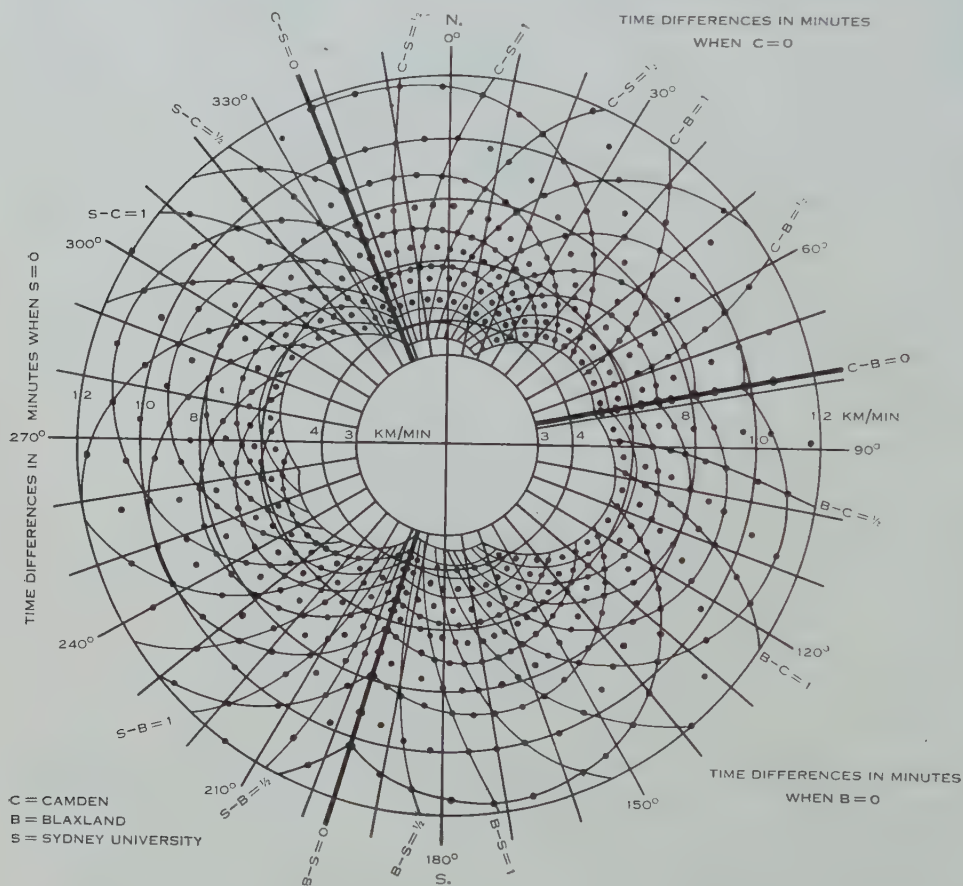


Fig. 4.—Chart showing limits of accuracy of deduction of directions and speeds from observational data.

For the speeds most frequently observed (5–10 km/min) the corresponding time differences are between 4 and 2 min. Tables have been computed to give the direction and speed for each possible pair of time differences at $\frac{1}{4}$ min intervals. Three sheets are used, one for zero time at each observing point. From these tables, it is easy to examine the change in direction and speed for a difference of $\frac{1}{4}$ or $\frac{1}{2}$ min in one of the time differences.

A plot of these points is shown in Figure 4 for the range of most commonly occurring velocities. The limits of accuracy are clearly shown for all directions

of movement. It will be seen that, where times are recorded to the nearest $\frac{1}{4}$ min, the lower limit of directional accuracy for normal velocities will be of the order of 5° .

(b) Variability

(i) *Daily Scatter*.—It has been mentioned previously (Munro 1950) that the observed directions and speeds may show considerable scatter in a single day. The limitations of the observing system have been discussed in an earlier section, but the variations are frequently greater than can be accounted for by such errors. Moreover, the directions may show regular drifts. Typical examples of the amount and variability of this scatter may be seen in Figure 5.

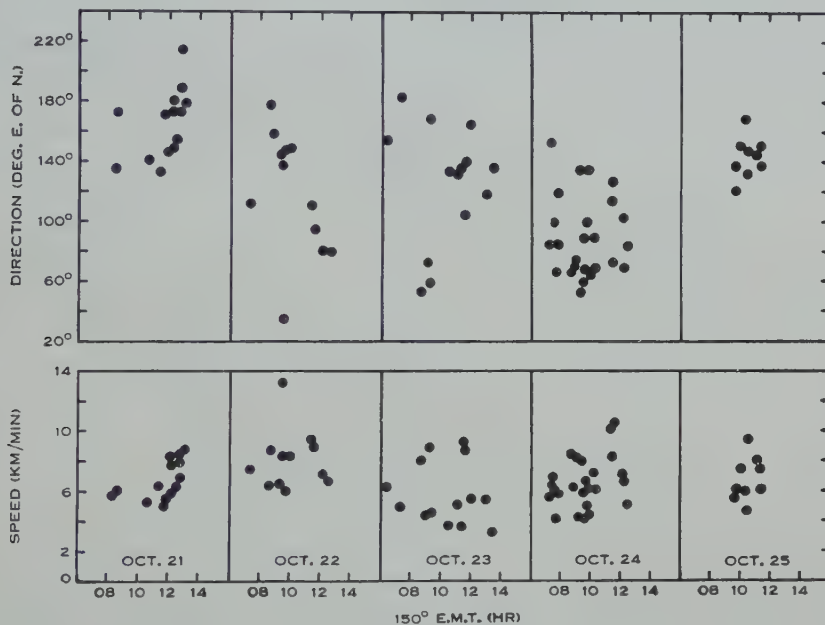


Fig. 5.—Day-to-day variations in direction and speed of movement, October 21–25, 1951.

The method of deducing directions assumes that each disturbance has a straight front. Curvature of the front could introduce an error in the deduced direction. In addition, any fluctuation local to one reflection point may bias the observation at that point. It has been apparent for some years that there are at least two scales of irregularities in the *F*-region reflections, one which produces fading periods of the order of seconds, and the other, of the type we are considering here, which gives fading periods of the order of minutes. This has been clearly confirmed by directional observations recorded by Bramley and Ross (1951). These smaller irregularities will be less liable to cause errors in deduced direction when complexities are observed than in the case of peaks or cross-overs, since they will affect the magnitude of a complexity without altering its position (or time of occurrence), but they may change the position (or time) of a retardation peak.

The tabulations will also tend to give some preferred directions owing to discontinuities in timing. For example, at speeds greater than 5 km/min the table as seen from Figure 4 gives no readings between 75° and 78° , and, omitting $\frac{1}{4}$ minutes, no readings between 70° and 78° ; and the gap increases with speed.

These limitations are taken into account in the statistical analyses which follow.

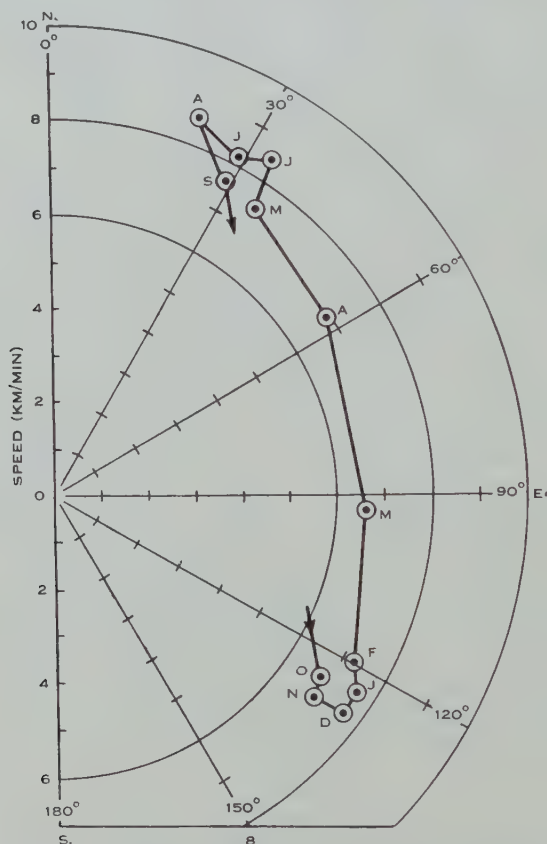


Fig. 6.—Seasonal variation of monthly median velocities.
Mean of 8 years: April 1948 to March 1956.

It should be borne in mind that the apparent direction, as observed in this way, will be the horizontal component of movement normal to the front of the disturbance.

(ii) *Day-to-day Variations.*—It has been mentioned earlier that the directions may vary considerably from day to day. This is particularly so at the equinoxes. Plots of all directions and corresponding speeds recorded each day for a period of 5 days are shown in Figure 5. It is obvious from inspection that the daily means of direction and speed vary appreciably.

Simultaneous observations at two points, Sydney and Perth (Price 1955), which have almost the same latitude but are separated by 3500 km, suggested

that the day-to-day variations in number of occurrences and direction of movement may have a longitude variation with an apparent lag in occurrence of approximately a day in this interval of 40° of longitude.

(c) *Seasonal Variation—Day-time*

(i) *Mean Directions and Speeds.*—The mean annual variations in monthly medians of apparent velocity of disturbances over a period of 1 year were shown in a previous paper (Munro 1950). Data now available cover a continuous period of 9 years, and the means of the monthly medians for 8 years are shown in Figure 6. The compact summer and winter groups appear very clearly, with an abrupt change in direction from September to October and a more gradual one from February to May; March and April being transitional months. Some additional features become more evident if the directions and speeds are considered separately and in more detail.

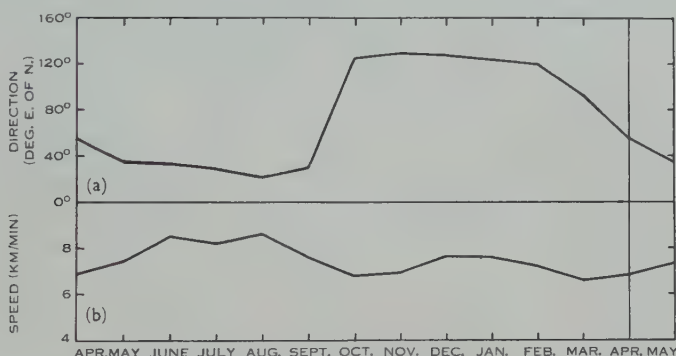


Fig. 7.—Monthly mean values of (a) direction of travel and (b) speed of horizontal movement, for the period April 1948 to May 1956.

In Figure 7 the data of Figure 6 are replotted to show in (a) the variation of the mean of the monthly medians of direction and (b) the corresponding variation in horizontal speed. The seasonal changes in direction are clearly shown in (a); and in (b) it will be seen that there is a definite maximum of speed in midwinter (June, July, and August) and a smaller maximum in midsummer. The minima occur in October and March, coinciding with the equinoctial changes of direction and also with the minima of occurrence shown in Figure 1 (c).

The general features of these annual variations have repeated from year to year with some secondary variations. The deviations from the mean curves are shown in Figures 8 (a) and 8 (b). In Figure 8 (a) it will be seen that, although there is some scatter of monthly values, which is greatest in October and February and greater in summer than in winter, no marked deviation from the mean direction persists for more than a few months. In Figure 8 (b) the year which shows the greatest abnormality is 1949–1950, where the speed was a record high value in June but fell to a record low value in July and continued low for the rest of the year, giving record lows for both minima. December 1948 shows a record summer maximum, so that we really have two high maxima in 1948

and 1949 followed by two low minima in 1949 and 1950. If we follow this clue and observe the total seasonal variations, that is, maxima and minima only, we get the result shown in Figure 9. It will be seen that the winter

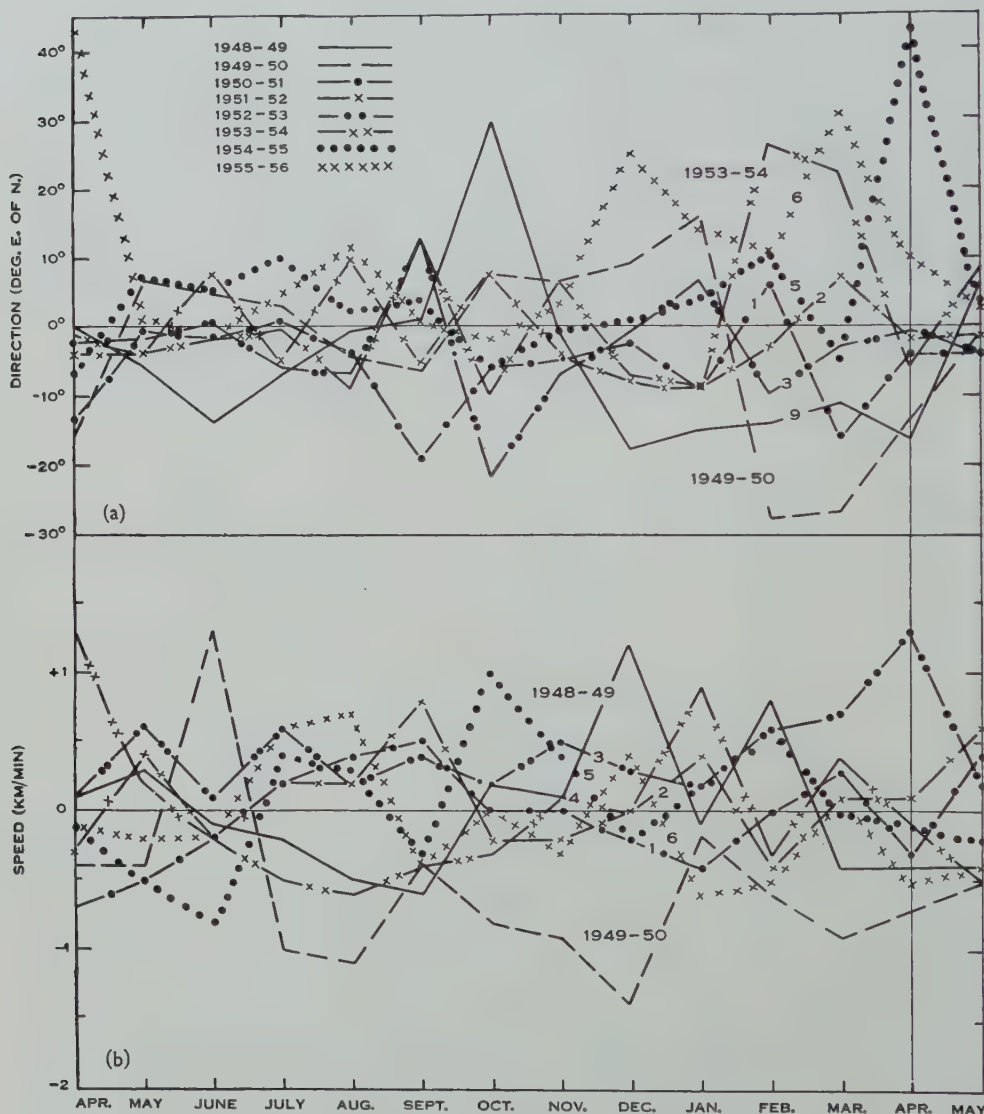


Fig. 8.—Deviations from mean of monthly medians of (a) directions, and (b) speeds. Period: April 1948 to May 1956.

maximum is greatest in 1949 and least in 1953, while the summer maximum is least in 1949–1950 and greatest in 1953. The variations for both equinoctial minima are similar, with trends roughly following the inverse of the summer maximum. 1953 happens to be the year of sunspot minimum, which may have

some significance, but rising values of winter maximum have not been maintained in 1956, which is a year of rapidly increasing sunspot numbers.

(ii) *Distribution of Directions and Speeds.*—(1) *Direction.*—It has been mentioned previously (Munro 1950) that the scatter of directions varies from month to month, being a minimum in midwinter. This has been confirmed in subsequent observations and merits closer examination. In Figure 10 are shown histograms of observed directions for three months—January (summer), March (equinoctial), and July (winter)—for each of three successive years—1951, 1952, and 1953. It will be seen that the pattern repeats its general form from year to year. This applies in all months, so the monthly values for 6 years have been lumped in Figure 11 to give the mean monthly distribution.

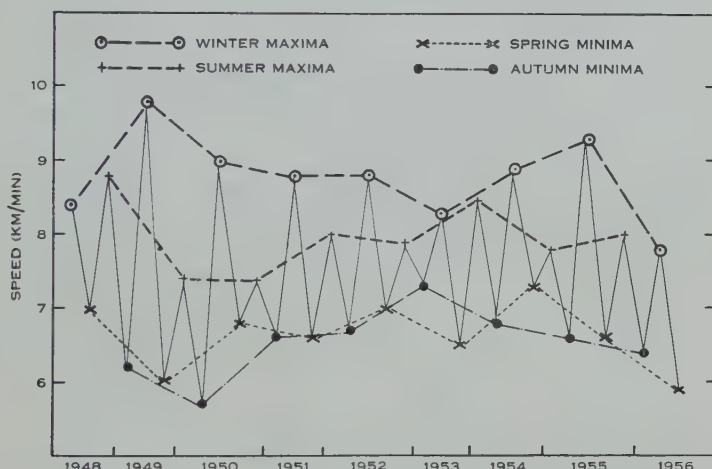


Fig. 9.—Variations in maxima and minima of monthly medians of speed.

This presentation suggests that there are two main groups of directions and that the seasonal drift is really due to a shift of emphasis from one group to the other. The winter group is the more compact and both groups are present in the transitional months, but the total spread for each month is approximately constant—about 180° .

The abrupt change from September to October and the more gradual one in March, April, and May is again apparent. The greater scatter in directions in the summer months, particularly December, is seen to be due to the indefinite peaks rather than to greater total spread.

(2) *Speeds.*—The monthly mean distribution of speeds is also consistent from year to year; in Figure 12 the monthly averages over 6 years are shown as histograms.

Very few values below 3 km/min are recorded, though there is no technical reason why they would not be observed. In all, 54 have been recorded in 7 years, ranging from 1.9 to 2.9. Values greater than 15 km/min are relatively scarce and those recorded are considered doubtful since the possible errors became great at speeds above 10 km/min, as can be seen in Figure 1.

The whole of the data for this period, comprising 13,642 values, are replotted on log-probability paper in Figure 13 with, for comparison, those for July, a month of maximum mean speed (1299 values) and those for March, a month of minimum mean speed (1029 values). It will be seen that in each case the points approximate closely to a straight line except for values above 90 per cent. As these are all for speeds greater than 10 km/min it is obvious that the deviation

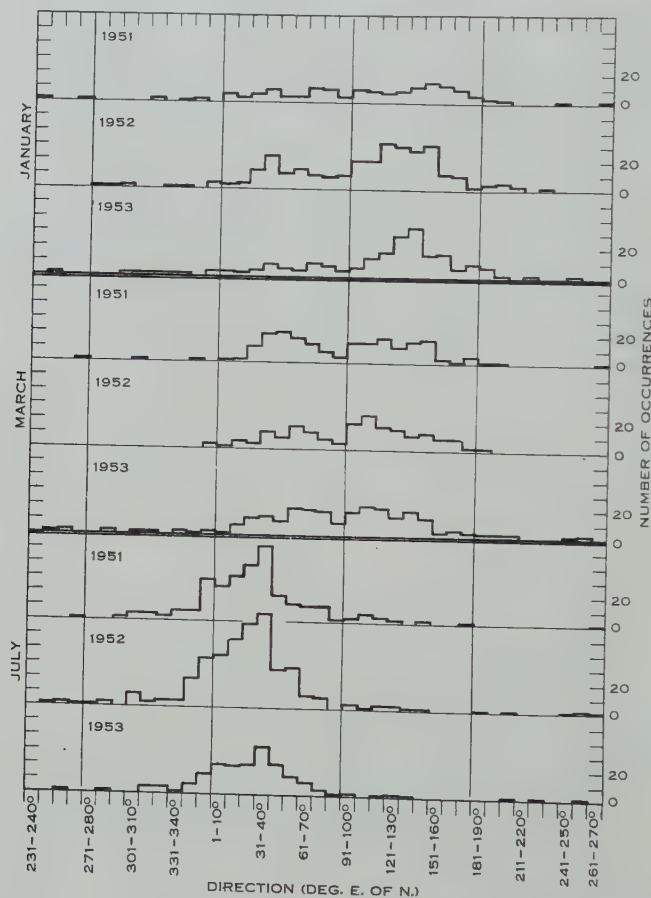


Fig. 10.—Distribution of directions during January, March, and July in the years 1951–1953.

is the result of biasing due to the method of computing speeds; it can be seen in Figure 4 that the possible errors are greatest at high speed and increase with each increment of time.

It appears therefore that the distribution is a normal Gaussian one.

(d) Diurnal Variation

It has not been possible to observe disturbances consistently over the full 24 hr. In the evening period of some 4 hr centred on sunset observable disturbances are rare, and interference from commercial radio stations, and

from atmospherics in summer, makes satisfactory recording difficult. In the early morning the penetration frequency is a minimum and was below 5.8 Mc/s for most of the years of observation; and it also seems to be a quiet period for disturbances. Recording has therefore been most consistent from 0730–1730 hr.

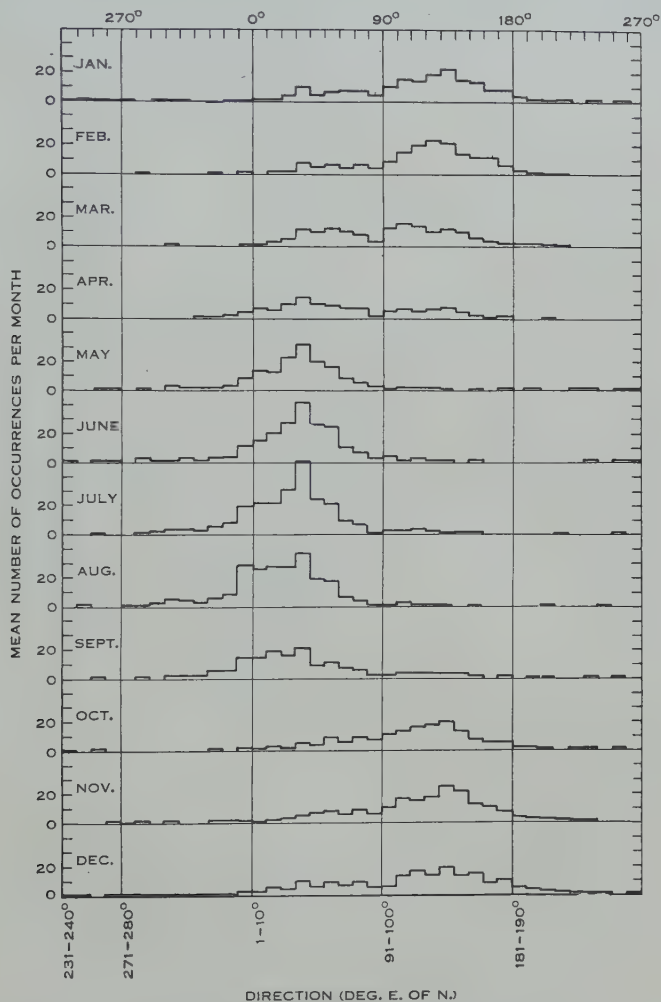


Fig. 11.—Distribution of directions for each month. Mean of years April 1950 to March 1956.

The night-time observations have been more in the nature of samplings during the most profitable periods but they have provided some significant results. These day and night observations will be considered separately.

(i) *Directions.* — (1) *Winter day-time.* — A definite drift in direction of movement during the period 0900 to 1500 hr in June was previously reported for the years 1950–1952 (Munro 1953*a*) and this has been repeated closely in subsequent years, the greatest deviation being in 1956. Even more consistent results

are evident in May, July, and August. The mean of the medians for each of these four winter months in the years 1950–1956 are shown in Figure 14. The obvious feature is a steady and consistent drift towards the north.

(2) *Winter night-time*.—Night observations were taken on 22 nights on a frequency of 2.28 Mc/s in June 1953, during the hours 2030–0200. Features observed were mostly branches from the main trace, which we interpret as a

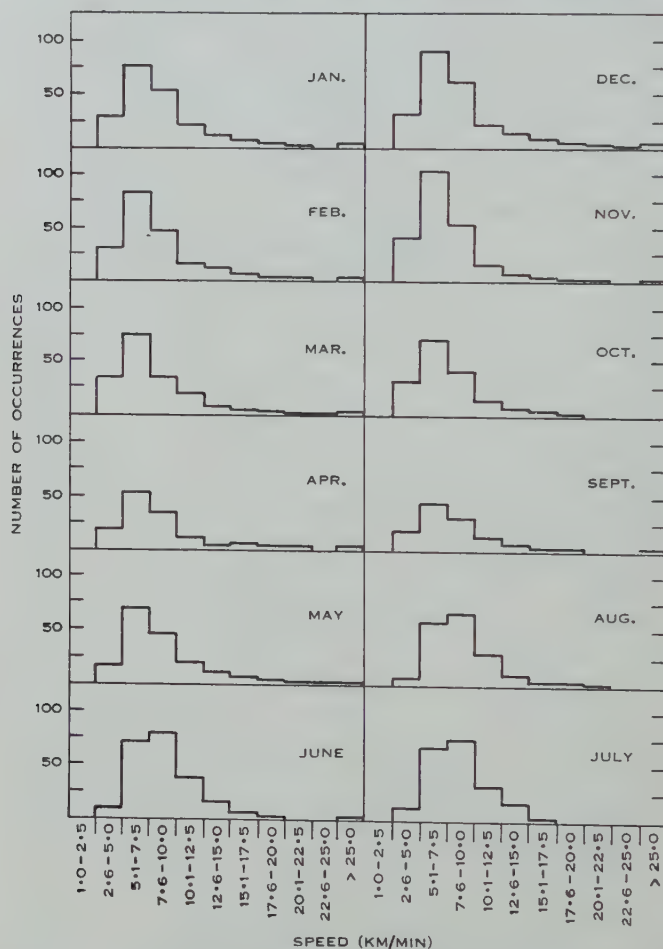


Fig. 12.—Distribution of speeds. Mean number of occurrences for each month of the year: period 1951–1956.

night-time version of a Y type anomaly (Munro 1953b), and there were some Z type anomalies and some sudden height changes; the results are shown in Figure 15. It will be seen that the directions are predominantly in the north-west-quadrant with a suggestion of a mean drift toward north about midnight.

(3) *Summer day-time*.—In summer over the years 1949–1954 no very definite diurnal variation was apparent, the hourly means being consistently close to 120° . In the 1955–56 summer, however, there were indications of a drift towards the

south in the afternoon, and in the 1956-57 summer the diurnal variation for November, December, and January was completely changed, showing a marked drift to the south and west. The contrast is shown in Figure 16 in which (a) shows the diurnal variation of mean hourly values for the months of November, December, and January in the years 1951-1954 and the maximum deviations from the mean values, while (b) shows the values in 1956-57.

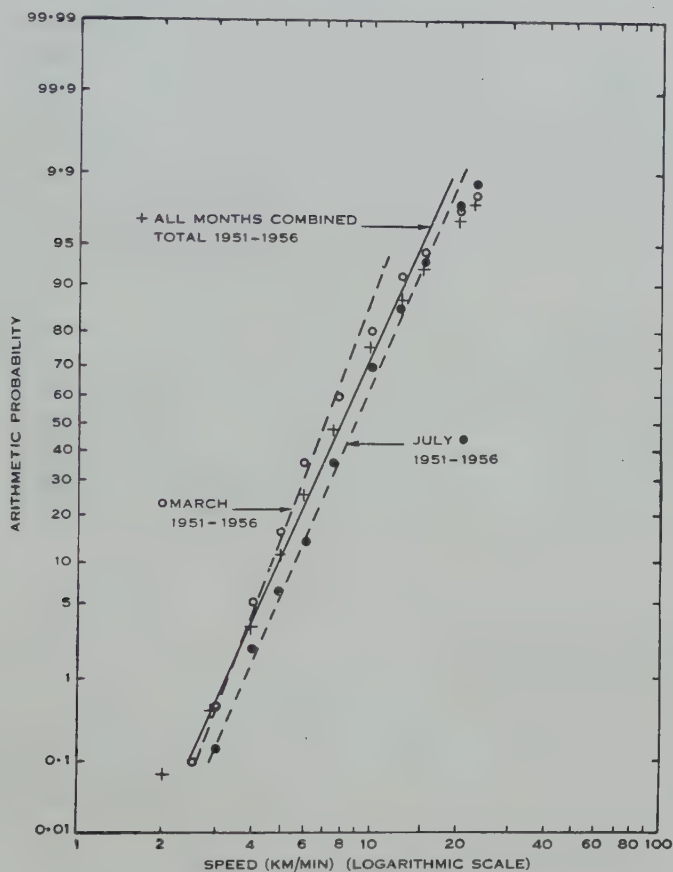


Fig. 13.—Distribution of speeds (probability plot).

Examination of directions recorded each day showed that during the earlier years there might be a definite drift of direction evident on a particular day but the direction of drift was not consistent from day to day. As an example the results for three consecutive days, December 5, 6, and 7, 1952 are shown in Figure 17, where it will be seen that there was an anticlockwise drift on the 5th, none on the 6th, and a clockwise drift on the 7th. There were also, in the earlier years, some directions recorded in the south-west section, mostly in the afternoon, but these were not sufficient to appreciably affect the monthly median and the number fell to a minimum in 1952-53.

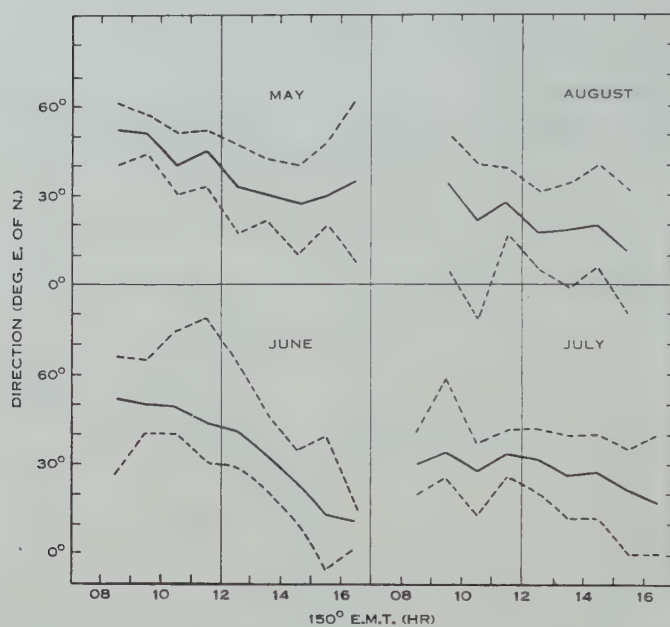


Fig. 14.—Mean day-time diurnal variation of direction in May, June, July, and August, 1950-1956. Broken-line curves show extreme values.

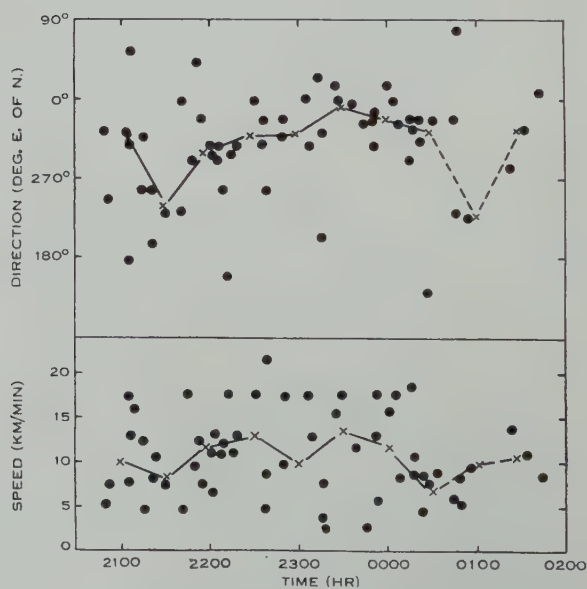


Fig. 15.—Diurnal variation of direction and speed, winter night (June 1953).

In November 1955 the midday reversals became more frequent and in December 1955 they showed a very marked increase. The difference between the morning and afternoon directions is emphasized in Figure 18: (a) shows the directional distributions before and after noon and it will be seen that the great

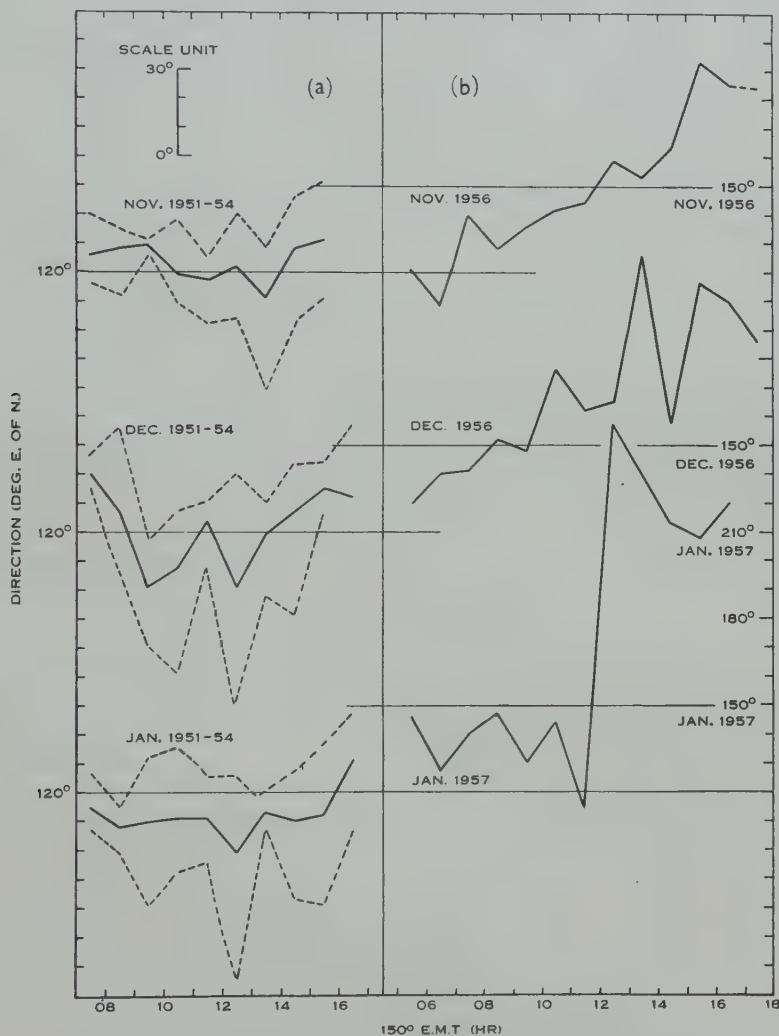


Fig. 16.—Diurnal variation in direction of movement: (a) Nov., Dec., Jan., 1951-1954 (broken line curves show extreme values); (b) Nov., Dec. 1956 and Jan. 1957.

majority of easterly directions occur in the morning and of westerly directions in the afternoon; (b) shows the time distribution of occurrences from 0 to 190° and from 191 to 260°, 190° being the approximate magnetic meridian direction. The easterly peak of occurrence is between 10 and 11 hr and the westerly peak between 13 and 14 hr.

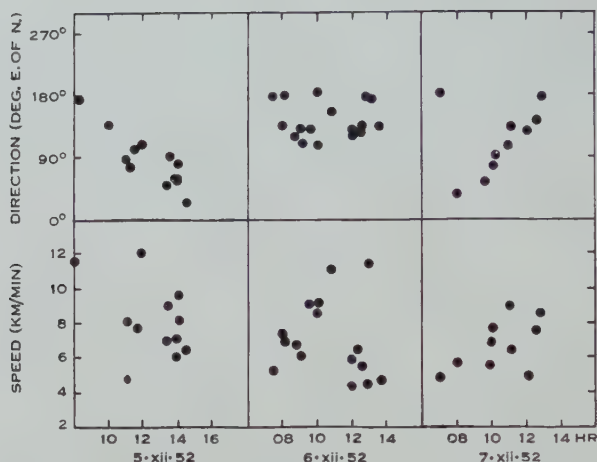


Fig. 17.—Diurnal drift of directions (Dec. 5, 6, 7, 1952).

(4) *Summer night-time*.—In the summer 1956–57, owing to the increased penetration frequencies and improved observational techniques, it has been possible to continue the observations on 5.8 Mc/s from 22 hr each night right through the morning and until 18 hr the next day.

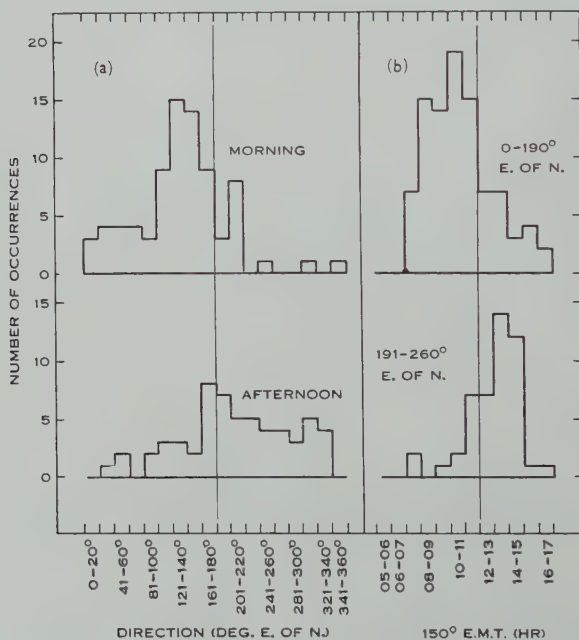


Fig. 18 (a).—Directional distribution of morning and afternoon observations.

Fig. 18 (b).—Time distribution of directions in south-east and south-west quadrants (Dec. 1955) in summer.

The results of these observations for January 1957 are shown in Figure 19. It will be seen that the east-west component appears to reverse again about midnight.

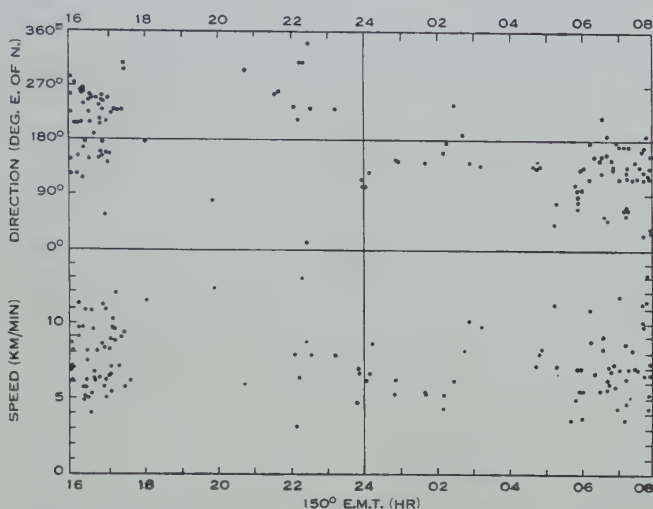


Fig. 19.—Summer night directions and speeds of travelling ionospheric disturbances, January 1957.

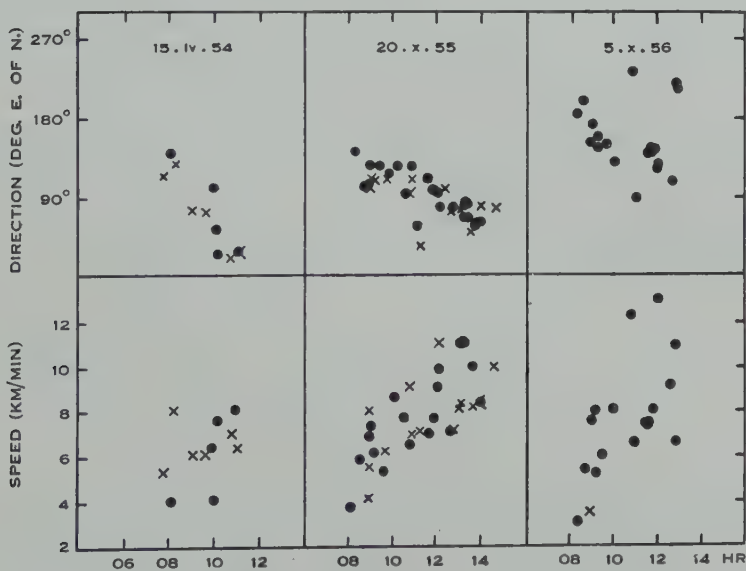


Fig. 20.—Diurnal drift of direction, equinoctial months (April 4, 1954; October 20, 1955; October 5, 1956).

(5) *Equinoctial day-time*.—In the equinoctial months, March, April, September, and October, directions generally show an irregular day-to-day fluctuation between summer and winter tendencies, with the new directions

progressing from a minority to a majority of days. The total change in direction is greater in the September-October equinox and seems to take place more rapidly. A clear mean diurnal variation would therefore seem unlikely and has not been found. There are, however, occasional days on which there is a

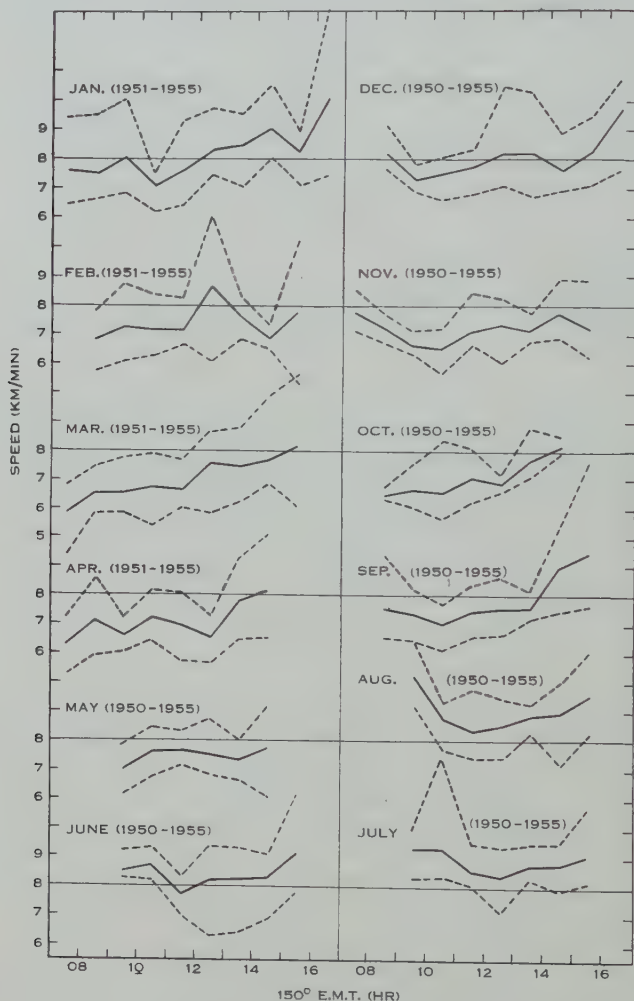


Fig. 21.—Diurnal variations of speed. Mean of medians for each hour.

consistent drift. Three such cases are shown in Figure 20 for the days April 15, 1954, October 20, 1955, and October 5, 1956. In each case the drift is in the same direction, i.e. anticlockwise.

(ii) *Speed*.—Diurnal variations of speed have been small, with little change from year to year. The mean of the medians for each hour of day-time observation during each month are shown in Figure 21, with the maximum deviations shown dotted. There does appear to be a consistent increase of the order of 30 per cent. in mean speed during the day, except in the winter.

The night values show little change from the day values for the corresponding seasons as can be seen in Figures 15 and 19.

It is of interest to note that, in the equinoctial days shown in Figure 20, where the direction has shown a definite drift from summer to winter directions the speed has also shown a corresponding drift from lower summer values to higher winter values. This is in agreement with the trend of the mean curves in Figure 21.

V. CONCLUSION

These observations have been all on a single frequency and therefore within a limited range of heights, which itself may have seasonal and sunspot variations. Examination of records has so far not shown any change of direction or speed with height. There do, however, appear to be variations in frequency of occurrence with height. Some indication of this from $h'f$ records is given in a paper by Heisler (1958) in this issue, and further observations are in progress to provide accurate information using two or three fixed frequencies which will be reflected at different heights. These latter observations will also give information on the slope of front of disturbances; as discussed by Munro and Heisler (1956), that is, the ratio of vertical to horizontal components of travel.

Improved techniques are also being used to give more complete information on diurnal changes. Studies of the dimensions and distance of travel of disturbances are also being continued. It is not proposed, therefore, to discuss in this paper the physical interpretation of the results.

The conclusions reached in Sections III and IV may be summarized as follows.

The mean number observed has been approximately six per day. The actual number recorded varies considerably from day to day. There is a marked maximum of occurrence about midday and possibly a smaller maximum at midnight. This maximum appears to have two peaks, one mainly in summer before noon and the other mainly in winter just after noon. Both peaks show clearly at the equinoxes. There is also a seasonal variation of occurrence with minima at the equinoxes.

Directions of travel have considerable scatter, but on most days a significant mean is deducible. This mean direction also shows some variation from day to day, but monthly median directions are significant and consistent from year to year.

There is a definite seasonal change of mean day-time direction from approximately 30° in winter to approximately 120° in summer, the change taking place more rapidly at the vernal equinox.

The winter diurnal variation has shown a consistent drift from approximately 60 to 20° between 0800 and 1600 hr, apparently continuing into the north-west quadrant at night.

In summer until recently there has been little variation during the day, but the very limited night observations suggest a change from the south-east quadrant by day to the south-west quadrant at night. During the last two summers,

however, there has been a definite change to the south-west quadrant near midday and a return change about midnight.

The variations in speed are not great. The scatter of values is found to have a normal probability distribution when examined on a monthly or yearly basis.

The mean monthly speed varies only between 7 and 9 km/min, with a main maximum in winter and a smaller one in summer, and approximately equal minima at the equinoxes. This variation is similar in form to that of occurrence. Ninety-eight per cent. of observed values fall between 3 and 20 km/min and the median value is 6 km/min.

The mean total day-time diurnal variation of speed is of the same order as the seasonal variation, showing a rise of approximately 2 km/min from 0800–1600 hr in the months from January to April and September and October.

Little correlation with other geophysical data has yet been found but the change in diurnal variation of direction during the last two summers suggests a connexion with sunspot activity.

VI. ACKNOWLEDGMENTS

This work is part of the investigations sponsored by the Radio Research Board. The recordings have been made at the Radio Research Laboratories of C.S.I.R.O. at Sydney and Camden and a field station at Blaxland. The analysis of results and preparation of material for the report have been carried out by the staff of the Sydney Laboratory located in the Electrical Engineering Department of the University of Sydney. The invaluable assistance of all members of the staff during the period of the observations, the general supervision by Dr. D. F. Martyn, the Chief Scientific Officer of the Radio Research Board, and the cooperation of the University of Sydney, and in particular Professor D. M. Myers, for provision of accommodation and facilities are gratefully acknowledged.

VII. REFERENCES

- BRAMLEY, E. N., and ROSS, W. (1951).—Measurements of the direction of arrival of short radio waves reflected at the ionosphere. *Proc. Roy. Soc. A* **207**: 251–67.
- HARVEY, J. A. (1955).—Movement of sporadic *E* ionization. *Aust. J. Phys.* **8**: 523–34.
- HEISLER, L. H. (1958).—Anomalies in ionosonde records due to travelling ionospheric disturbances. *Aust. J. Phys.* **11**: 79.
- MUNRO, G. H. (1950).—Travelling disturbances in the ionosphere. *Proc. Roy. Soc. A* **202**: 208–23.
- MUNRO, G. H. (1953a).—Travelling disturbances in the ionosphere: diurnal variation of direction. *Nature* **171**: 693.
- MUNRO, G. H. (1953b).—Reflexions from irregularities in the ionosphere. *Proc. Roy. Soc. A* **219**: 447–63.
- MUNRO, G. H., and HEISLER, L. H. (1956).—Divergence of radio rays in the ionosphere. *Aust. J. Phys.* **9**: 359–72.
- PRICE, R. E. (1955).—Travelling disturbances in the ionosphere. In “The Physics of the Ionosphere,” pp. 181–90. (Phys. Soc.: London.)

THE 1956 PHOENICID METEOR SHOWER

By A. A. WEISS*

[*Manuscript received November 18, 1957*]

Summary

From radio observation of this shower at Adelaide the radiant coordinates are estimated to be 15 ± 2 , -55 ± 3 . The radio record was obtained when the Earth was some 6 hr from the centre of the stream. The radio rate of 30/hr measured on an equipment of high sensitivity is much lower than expected from the visual rates of from 20 to 100/hr reported from 1 to 9 hr later. Echo duration and amplitude are smaller than would be expected from the visual brightness of these meteors. The low radio rate and lack of bright radio meteors could be due to observation on the fringe of the stream or to low ionizing efficiency of slow meteors.

I. INTRODUCTION

Visual observations of a new meteor shower, active on the night of December 5, 1956, have been reported by Ridley (1957) and by Shain (1957), who determined radiants in the constellation Phoenix, at 15 , -45 and 15 , -58 respectively. Orbital elements, computed by Ridley, are similar to the elements of Comet 1819 IV Blanpain as calculated from one apparition only.

This shower was recorded on the 67 Mc/s narrow-beam radar equipment at Adelaide during a routine survey of meteor activity in the southern hemisphere. Since the equipment was described by Weiss (1955a) the transmitter power has been increased considerably. At the time of the shower only one recording channel was in operation, and echoes were received only from the S. aerial, whose beam axis is directed at azimuth 14° N. of E., elevation 9° .

II. RADIANT AND ORBIT

A conventional range-time plot of all echoes received from 19.40 to 22.20 hr L.T. is given in Figure 1. As only one aerial was in use, the full potentiality of the equipment for accurate determination of shower radiants could not be realized and recourse to the envelope-fitting method was necessary. Fortunately, for a radiant located so close to the south celestial pole the shape of the range-time envelope is quite sensitive to the declination of the radiant, and the declination can be measured accurately. The error in Right Ascension inherent in the process of fitting a range-time envelope to the observed echoes is small. Range-time envelopes have been calculated assuming a mean meteor height of 90 km. As Phoenicid meteors overtake the Earth from behind, the velocity of entry into the Earth's atmosphere (geocentric velocity) must be low and the mean height should accordingly be somewhat lower than 90 km. However, the error in radiant coordinates introduced by uncertainty in the height is small.

* Division of Radiophysics, C.S.I.R.O., at Department of Physics, University of Adelaide.

The radiant coordinates are estimated to be 15 ± 2 , -55 ± 3 , near the star ξ Phoenix. The fitted range-time envelope in Figure 1 indicates that the radiant area is small.

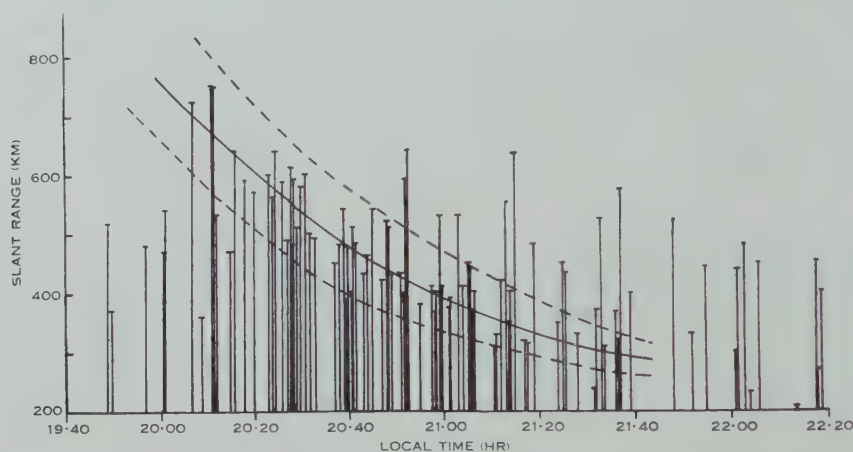


Fig. 1.—Range-time plot of Phoenicid echoes. The fitted envelope is for radiant coordinates 15, -55 .

Orbital elements have been computed for this radiant, assuming (a) parabolic velocity and (b) a period of 5.1 years as suggested by the general similarity

TABLE I
ORBITAL ELEMENTS OF THE PHOENICIDS

Element	Assumed Period	
	∞	5.1 Years
Apparent radiant ..	15, -55	15, -55
Corrected radiant ..	15, -56	15, -58
Apparent elongation ..	125.1°	123.6°
True elongation	160.7°	164.1°
Ω	73.4°	73.4°
ω	0.1°	0.3°
i	19.3°	15.9°
e	1.0	0.667
q	0.985	0.985
Heliocentric velocity ..	42.5 km/sec	38.8 km/sec
Apparent velocity ..	17.1	12.7
Geocentric velocity ..	20.4	16.9

of the parabolic elements with those of Comet 1819 IV Blanpain. These new elements are given in Table 1, along with some data on velocities. They do not agree with the elements of Comet 1819 IV Blanpain as well as the elements already

computed by Ridley (1957) for the more northerly radiant. With the exception of the eccentricity, it is clear that the general description of the meteor orbit is not much affected by our lack of precise knowledge of the radiant and velocity.

III. RADAR AND VISUAL RATES

The absolute sensitivity of the radio equipment has not been determined, but early in December 1956 the sporadic echo rate was about 600 per day. This far exceeds the visual rate, but, as the character of the diurnal variation at high radio rates appears to differ from the visual diurnal variation, rate comparisons of this type are not a useful index of equipment sensitivity. A more reliable index is provided by the permanent showers.

Radar rates, measured with the Adelaide equipment at high sensitivity, are compared with corresponding visual rates for three showers in Table 2. Visual rates are corrected to a radiant at the zenith. The δ -Aquarid radiant passes through the aerial collecting zone (inclined at an angle of 14° to the north-south

TABLE 2
ZENITHAL RATES OF METEOR SHOWERS

Shower	Radio Rate	Visual Rate	Authority for Visual Rate
Phoenicids	30	20—100 > 60	Ridley (1957) Shain (1957)
δ -Aquarids	150	20	} Lovell (1954)
Geminids	120	20—60	

meridian) at roughly the same zenith angle, 20° , as the Phoenicid radiant and the geometry of detection for the two radiants will be similar. This is borne out by comparison of the theoretical range-time envelopes. As the two radiants transit so close to the zenith, echo rates have not been corrected. Consideration of the geometry of detection of the Geminid radiant shows that the zenithal correction factor is approximately \sec^s-1z , where z is the zenith angle of the radiant at detection (here 76°) and s is the mass-distribution parameter ($s \sim 1.5$ for the Geminids).

Despite the heterogeneous nature of the rates for the permanent showers, there can be no doubt that the echo rate for the Phoenicids, in relation to the visual rates, is surprisingly low. The significance of this low rate is discussed below.

IV. ECHO DURATION AND AMPLITUDE

Echo duration and amplitude were not measured, but some information on these characteristics may be obtained in the following way. At the slow film speed used (12 cm/hr) spot size and intensity are determined jointly by the echo amplitude, which was not voltage-limited, and the echo duration. As large echo amplitude is usually associated with long echo duration, spot size and intensity furnish a rough guide to the echoing area of the meteor trail. Spot intensities

have been divided subjectively into four classes and the number-distribution of 59 echoes falling within the fitted range-time envelope is listed in Table 3. This distribution for Phoenicid meteors does not differ significantly from the distribution for an equal number of sporadic meteors taken on either side of the shower.

The inference is that at the time of the Adelaide observations, which extended from 11 to 12.30 hr U.T., the Phoenicids included very few bright radio meteors. This is remarkable in view of the visual observations made very little later. Thus Shain, observing at Sydney, reports visual apparent magnitudes estimated to be about -2 ; these observations, made from 13 to 13.30 hr U.T., were limited by cloud. Still later, from 16.40 to about 22 hr U.T., South African visual observers reported many meteors of fireball magnitude, with maximum activity occurring round 19 hr U.T. Unusual activity was also reported from New Zealand about 10 hr before the onset of the shower in South Africa. No unusual visual objects were reported at Adelaide, but the sky was almost completely obscured by cloud.

TABLE 3
COMPARISON OF PHOENICID AND SPORADIC ECHO
INTENSITIES

Echo Classification			Phoenicids	Sporadics
1	Weak	14	16
2	18	13
3	18	21
4	Strong	9	9

It is possible that the low radio rate and the deficiency of bright radio meteors resulted from detection at Adelaide whilst the Earth was still on the fringe of the stream. However, the visual record suggests a marked concentration of massive meteor particles to the centre of the stream, without any corresponding concentration of meteor density; and the radiant passed out of the Adelaide collection zone less than an hour before bright visual meteors were reported from Sydney, some 6 hr before the peak activity. Under these circumstances there is an alternative explanation of the Adelaide observations which merits consideration.

The shortest echo duration which can be resolved is limited by film speed and spot size, and is a little less than 10 sec. None of the Phoenicid echoes had durations exceeding this, and the great majority of echo durations were considerably less.

The relation between visual brightness and line density α of electrons/cm in the meteor trail is usually taken to be

$$\log_{10}\alpha = 14.0 - 0.4M_v, \quad \dots\dots\dots (1)$$

The constant is determined by the condition that $M_v = 5$ corresponds to $\alpha = 10^{12}$ (see e.g. Browne *et al.* 1956). Substitution of Shain's estimated visual brightness

of $M_v = -2$ in (1) gives $\alpha = 6 \times 10^{14}$. Trails with electron densities as high as this give persistent echoes whose duration (Kaiser 1953) is

$$\tau = 1.124 \times 10^{-12} (\lambda^2 / 16\pi^2 D) \alpha. \quad \dots\dots\dots (2)$$

With $\lambda = 448$ cm and $D = 4 \times 10^4$ cm²/sec, corresponding to a mean height of 90 km (Weiss 1955*b*), $\tau = 21$ sec. As the Phoenicid meteor velocity is low, the mean height is probably lower than 90 km, which implies a lower mean value for D and hence longer durations. In any case, the height distribution of the shower meteors would also provide some lower values of D . The absence of echoes with resolvable durations would therefore suggest that the visual brightness of the meteors detected at Adelaide was considerably less than that corresponding to $M_v = -2$.

However, relation (1) takes no account of a possible dependence of ionizing efficiency on velocity. The evidence in favour of a strong increase in ionizing efficiency with increasing velocity has recently been reviewed by Weiss (1957). If the expression given by Hawkins (1956) is adopted, (1) must be replaced by

$$\log_{10} \alpha = 11.52 + 1.56 \log_{10} v - 0.40 M_v. \quad \dots\dots\dots (3)$$

with v in km/sec. Taking as an upper limit for the geocentric velocity $v = 20$ km/sec corresponding to a parabolic orbit, (2) and (3) set an upper limit of $\tau = 8$ sec for visual brightness $M_v = -2$. Even after allowing for meteors below the mean height of 90 km, the absence of resolvable echo durations is now consistent with the visual observations made at about the same time. Low ionizing efficiency would also depress the radio echo rate relative to the visual rate.

Even if the low velocity of the Phoenicid meteors does not afford a complete explanation of the radio observations, it may well be a contributing factor, whose importance can be assessed by combined radar/visual observations on subsequent returns of this stream.

V. REFERENCES

- BROWNE, I. C., BULLOUGH, K., EVANS, S., and KAISER, T. R. (1956).—*Proc. Phys. Soc. Lond.*, B **69** : 83.
 HAWKINS, G. S. (1956).—*Astrophys. J.* **124** : 311.
 KAISER, T. R. (1953).—*Advanc. Phys.* **2** : 495.
 LOVELL, A. C. B. (1954).—“*Meteor Astronomy*.” (Clarendon Press : Oxford.)
 RIDLEY, H. B. (1957).—*Circ. Brit. Astr. Ass.*, No. 382.
 SHAIN, C. A. (1957).—*Observatory* **77** : 27.
 WEISS, A. A. (1955*a*).—*Aust. J. Phys.* **8** : 148.
 WEISS, A. A. (1955*b*).—*Aust. J. Phys.* **8** : 279.
 WEISS, A. A. (1957).—*Aust. J. Phys.* **10** : 397.

ELECTROMAGNETIC PROPAGATION IN AN ALMOST HOMOGENEOUS MEDIUM

By V. W. BOLIE*

[*Manuscript received September 30, 1957*]

Summary

This paper concerns the development from Maxwell's electromagnetic equations of an equation of propagation in an almost homogeneous medium. The equation is applied to the problem of determining the secondary wave produced by an isolated Gaussian-shaped perturbation in the refractive index. An exact solution is obtained for points located on the axis of symmetry parallel to the direction of propagation of the incident primary wave. An approximate solution for points remote from the anomaly is obtained and its validity is compared with the more restricted exact solution. An interesting limit process is encountered in the derivation of the formula for the scattering cross section of the refractive index perturbation.

I. INTRODUCTION

In the field of radio communication there is currently a great deal of interest in the scattering of high-frequency waves by a turbulent atmosphere and the trans-horizon propagation of measurable signal strengths. Experiments with propagation of microwaves beyond the radio horizon show signal characteristics which cannot be explained in terms of free-space propagation, horizon diffraction, or mode theory of tropospheric ducts. In 1950 Booker and Gordon† proposed a scattering theory based on the random space-dependence of the dielectric constant due to atmospheric turbulence. Although this theory was promising in some respects, the detail of the scattering mechanism was restricted to a random array of dipole scatterers. A number of other theoretical papers and a large amount of experimental data have recently appeared in the literature, none covering extensions of the simple dipole theory.

The purpose of this paper is to develop fundamental equations for the propagation and scatter of electromagnetic energy in a nearly transparent medium, and to apply these equations to the problem of scattering by an isolated, Gaussian-shaped perturbation in the refractive index. The resultant solution for the scattered field provides a means of determining the scattering caused by a refractive anomaly of arbitrary size.

* Collins Radio Company, Cedar Rapids, Iowa; present address: Iowa State College, Ames, Iowa, U.S.A.

† BOOKER, H. F., and GORDON, W. E. (1950).—A theory of radio scattering in the troposphere. *Proc. Inst. Radio Engrs.*, N.Y. 38: 401–12.

II. PROPAGATION IN A NEARLY HOMOGENEOUS MEDIUM

The propagation of electromagnetic energy is, in general, described by the partial differential equations of Maxwell. For a linear, isotropic, charge-free medium of zero conductivity, Maxwell's equations can be used to show that the electric intensity vector $E(x, y, z, t)$ is governed by the equation

$$\nabla^2 E - \nabla(\nabla \cdot E) = \mu \epsilon (\partial^2 E / \partial t^2), \quad \dots \dots \dots (1)$$

where μ and ϵ represent the permeability and permittivity of the medium. The free-space values for μ and ϵ are $\mu_0 = 4\pi/10^7$ H/m and $\epsilon_0 = 10^{-9}/36\pi$ F/m respectively.

For a stationary, almost homogeneous medium, the refractive index $n = \sqrt{(\mu\epsilon/\mu_0\epsilon_0)}$ may be expressed as

$$n = n_1[1 + p(x, y, z)], \quad \dots \dots \dots (2)$$

where n_1 is a dimensionless constant slightly greater than unity, and where $|p(x, y, z)| \ll 1$. For a medium like the atmosphere, the permeability $\mu = \mu_1$ may be assumed constant, so that the permittivity is approximated by the equation

$$\epsilon = \epsilon_1[1 + 2p(x, y, z)], \quad \dots \dots \dots (3)$$

where $\epsilon_1 = n_1^2 \mu_0 \epsilon_0 / \mu_1$.

A convenient substitution for the divergence of E in equation (1) can be obtained by substituting equation (3) into the Maxwell relation $\nabla \cdot (\epsilon E) = 0$. After suitable manipulation of vector identities, followed by a logarithmic expansion, the approximation

$$\nabla \cdot E = -2E \cdot \nabla p \quad \dots \dots \dots (4)$$

is obtained. The substitution of equation (4) into equation (1) gives,

$$\nabla^2 E + 2\nabla(E \cdot \nabla p) = \mu_1 \epsilon_1 (1 + 2p) \partial^2 E / \partial t^2. \quad \dots \dots \dots (5)$$

Assuming the sinusoidal time-dependence, $E = E' \sin \omega t$, gives,

$$\nabla^2 E' + \nabla(E' \cdot \nabla p) + (4\pi^2/\lambda_1^2)(1 + 2p)E' = 0, \quad \dots \dots \dots (6)$$

in which the conventional notation $\lambda_1 = 2\pi/\omega \sqrt{(\mu_1 \epsilon_1)}$ has been used.

It is convenient at this point to assume the total field E' as being composed of a weak scattered field \tilde{E} , and a dominant homogeneous field \bar{E} which satisfies the simple propagation law $\nabla^2 \bar{E} + (4\pi^2/\lambda_1^2)\bar{E} = 0$. Making this assumption, and neglecting the effects of secondary scattering of the scattered field, gives

$$\nabla^2 \tilde{E} + (4\pi^2/\lambda_1^2) \tilde{E} = -(8\pi^2/\lambda_1^2) p \bar{E} - 2\nabla(\bar{E} \cdot \nabla p), \quad \dots \dots \dots (7)$$

which shows how the direct homogeneous field \bar{E} produces sources for the scattered field \tilde{E} .

In the foregoing propagation equation for the scattered field \tilde{E} it is seen that the scattering excitation is represented by two separate terms on the right side of the equation. The scattering associated with the term $(8\pi^2/\lambda_1^2) p \bar{E}$ has been examined on intuitive grounds by other investigators. The existence of the term $2\nabla(\bar{E} \cdot \nabla p)$ has been neglected or ignored in most of the literature

on scattering theory, even though its effect may be appreciable under certain conditions. If $p(x, y, z)$ is an arbitrary function, then ∇p is a vector of arbitrary magnitude and direction, and $\nabla(\vec{E} \cdot \nabla p)$ will also be arbitrary in magnitude and direction. This is in contrast with the vector $(8\pi^2/\lambda_1^2)p\vec{E}$, which is always parallel to the incident direct-wave field vector \vec{E} .

III. SCATTERING FROM A NEARLY TRANSPARENT ANOMALY

In order to examine in detail the mechanism of radiation scattering in a nearly homogeneous medium, it is convenient to consider the effect of an isolated perturbation in the refractive index. While the incident-direct field may be taken as a uniform plane wave without appreciable loss of generality, the refractive-index perturbation should be three-dimensional and continuous. Such a model is well approximated by assuming the fractional variation $p(x, y, z)$ of the refractive index has its maximum value at the origin and decreases with radial distance from the origin in a Gaussian manner.

This is expressed mathematically as $p = p_0 \exp(-r^2/s^2)$, where p_0 is the maximum value of p , s represents the "anomaly radius", and $r = \sqrt{(x^2 + y^2 + z^2)}$. The incident-direct field may be taken as a plane wave having its electric intensity parallel to the x -axis and travelling in the z -direction. Thus $\vec{E} = \hat{u}_1 E_0 \exp(-j2\pi z/\lambda_1) + \hat{u}_2 \cdot 0 + \hat{u}_3 \cdot 0$, where E_0 is the necessary amplitude constant, where j is the unit imaginary number, and where $\hat{u}_1, \hat{u}_2, \hat{u}_3$ represent unit vectors in the x, y, z directions.

Substituting the above assumptions into equation (7), and making use of the well-known retarded potential theory, shows that the components of the scattered-field vector $\mathbf{E} = u_1 \mathbf{E}_1 + u_2 \mathbf{E}_2 + u_3 \mathbf{E}_3$ are given by the integrals

$$\mathbf{E}_1 = \frac{p_0 E_0}{4\pi} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \frac{4}{\rho} \left[\frac{2\pi^2}{\lambda_1^2} + \frac{2\xi^2 - s^2}{s^4} \right] h(\xi, \eta, \zeta) d\xi d\eta d\zeta, \quad \dots \quad (8)$$

$$\mathbf{E}_2 = \frac{p_0 E_0}{4\pi} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \frac{8}{\rho} \frac{\xi\eta}{s^4} h(\xi, \eta, \zeta) d\xi d\eta d\zeta, \quad \dots \quad (9)$$

$$\mathbf{E}_3 = \frac{p_0 E_0}{4\pi} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \frac{8}{\rho} \left[\frac{\xi\zeta}{s^4} + j \frac{\pi}{\lambda_1} \frac{\xi}{s^2} \right] h(\xi, \eta, \zeta) d\xi d\eta d\zeta, \quad \dots \quad (10)$$

where

$$h(\xi, \eta, \zeta) = \exp[-\xi^2 + \eta^2 + \zeta^2]/s^2 - j2\pi(\zeta + \rho)/\lambda_1]$$

and

$$\rho^2 = (x - \xi)^2 + (y - \eta)^2 + (z - \zeta)^2.$$

These expressions are too complicated to be readily evaluated without the aid of automatic computing equipment. Fortunately, however, there are two important cases of practical interest for which the integrals can be simplified

in terms of known functions. One case is when the field point (x, y, z) is on the z -axis, and the other case is when $r \gg s$ so that the integrands are negligibly small except where the value of ρ in the denominators may be assumed equal to r and the value of ρ in the exponents is well approximated by its projection in the direction of the field point.

IV. EXACT SOLUTION IN THE INCIDENT-WAVE DIRECTION

At points far removed from the scattering anomaly the scattered field integrals given in equations (8), (9), and (10) can be simplified and evaluated in terms of elementary functions. The resultant solutions, however, can only be considered as approximations which are asymptotic to the exact solutions. The approximations will degenerate as the field point approaches the scattering anomaly. In order to estimate the validity of the approximations it is convenient to determine the exact solutions under conditions sufficiently restricted to permit the evaluation of the complicated integrals. Such conditions result from assuming the field point is on the z -axis.

If the field point (x, y, z) is on the z -axis, then $x=y=0$, and the integrands in equations (9) and (10) are odd in ξ , with the result that \mathbf{E}_2 and \mathbf{E}_3 are both zero. This result can also be reasoned from the symmetry of the physical problem. In order to simplify the integral for \mathbf{E}_1 , on the z -axis, it is convenient to introduce the translated polar coordinate notation $\xi = \rho \sin \alpha \cos \beta$, $\eta = \rho \sin \alpha \sin \beta$, $\zeta = z + \rho \cos \alpha$, and to replace the volume element $d\xi d\eta d\zeta$ by $\rho^2 \sin \alpha d\alpha d\beta d\rho$. Under these conditions the integrand in equation (8) is independent of β , permitting immediate integration with respect to that variable. Integration with respect to α is easily done by replacing $\cos \alpha$ by a dummy variable. The final integration with respect to ρ gives

$$\mathbf{E}_1 = [C_1 I_1 - C_2 I_2 + C_3] p_0 E_0 \frac{\sqrt{\pi}}{2} \frac{s}{z} \left(\frac{2\pi s}{\lambda_1} \right)^2 \exp(-j2\pi z/\lambda_1), \quad \dots \quad (11)$$

where

$$C_1 = \frac{1}{(1 + j\pi s^2/\lambda_1 z)} - \frac{j\lambda_1/2\pi z}{(1 + j\pi s^2/\lambda_1 z)^2} - \frac{\lambda_1^2/4\pi^2 z^2}{(1 + j\pi s^2/\lambda_1 z)^3}, \quad \dots \quad (12)$$

$$C_2 = \frac{1}{(1 + j\pi s^2/\lambda_1 z)} - \frac{j\lambda_1/2\pi z}{(1 + j\pi s^2/\lambda_1 z)^2} - \frac{\lambda_1^2/4\pi^2 z^2}{(1 + j\pi s^2/\lambda_1 z)^3}, \quad \dots \quad (13)$$

$$C_3 = \frac{\lambda_1^2/2\pi^2 z}{(1 + j\pi s^2/\lambda_1 z)^2} \frac{\exp(-z^2/s^2)}{s\sqrt{\pi}}, \quad \dots \quad (14)$$

and where

$$I_1 = \frac{1}{2} + \frac{1}{s\sqrt{\pi}} \int_0^z \exp(-u^2/s^2) du, \quad \dots \quad (15)$$

$$I_2 = \frac{\exp(j4\pi z/\lambda_1)}{s\sqrt{\pi}} \int_z^\infty \exp(-u^2/s^2 - j4\pi u/\lambda_1) du, \quad \dots \quad (16)$$

The quantities C_1 , C_2 , and C_3 are complex numbers which are readily evaluated in terms of s , λ_1 , and z . The integral I_1 is easily interpreted in terms of the error function of well-known statistical theory. The integral I_2 can be represented in terms of the error function of a complex variable.

Equation (11) represents the exact solution for the scattered field on the z -axis. The formula is valid for all values of z , including points at or near the origin, where the centre of the scattering anomaly is located. When the anomaly radius is large compared to one wavelength, the above solution can be further simplified for points near the origin. Thus, if $s \gg \lambda_1$, and $-s \leq z \leq 3s$, it can be shown that the quantity $C_1 I_1 - C_2 I_2 + C_3$ is well approximated by the value $C_1 I_1 - C_2 I_2 + C_3 \cong (-j\lambda_1 z / 2\pi s^2)[1 + \operatorname{erf}(z)]$. Substituting this expression into equation (11) gives

$$E_1 = -jp_0 E_0 [1 + \operatorname{erf}(z)] \frac{\sqrt{\pi}}{2} \frac{2\pi s}{\lambda_1} \exp(-j2\pi z / \lambda_1), \quad \dots \quad (17)$$

which is valid for $-s \leq z \leq 3s$ when $s \gg \lambda_1$. As might be anticipated on intuitive grounds, the magnitude of the scattered field intensity increases with z in the vicinity of the origin. This effect is illustrated in Figure 1, where the space factor

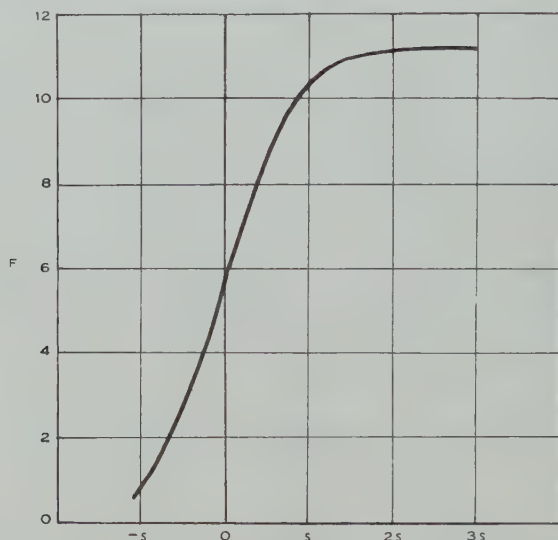


Fig. 1.—Space factor for the scattered field. On the z -axis the scattered field is given by $|E/E_0| = p_0 s F / \lambda_1$. When $s \gg \lambda_1$ and

$$-s \leq z \leq 3s, \quad F = \pi \sqrt{\pi} \{1 + (2/s \sqrt{\pi}) \int_0^z \exp(-u^2/s^2) du\}.$$

$F = \pi \sqrt{\pi} [1 + \operatorname{erf}(z)]$ is plotted as a function of z . It is apparent from the graph that the scattered field intensity does not attain its greatest magnitude until the incident wave has passed through the greater part of the refractive region.

Equation (11) is easily simplified for large values of z . It is readily shown that when $s > \lambda_1$ and $2s < z < \infty$, the value of I_1 is nearly unity, the magnitudes of C_3 and $C_2 I_2$ are negligible, while the expression for C_1 simplifies so that $C_1 I_1 - C_2 I_2 + C_3 \cong [1 + j\pi s^2 / \lambda_1 z]^{-1}$. Equation (11) then reduces to

$$E_1 = p_0 E_0 \frac{\sqrt{\pi}}{2} \left(\frac{2\pi s}{\lambda_1} \right)^2 \frac{\exp(-j2\pi z / \lambda_1)}{z/s + j\pi s / \lambda_1}, \quad \dots \quad (18)$$

which is valid when $s > \lambda_1$ and $z > 2s$. It is seen from this expression that as z increases in the positive direction the amplitude of the scattered field eventually varies inversely with distance. Thus, when $z \gg \pi s^2/\lambda_1$, the above formula becomes

$$E_1 = p_0 E_0 \frac{\sqrt{\pi}}{2} \left(\frac{s}{z} \right) \left(\frac{2\pi s}{\lambda_1} \right)^2 \exp(-j2\pi z/\lambda_1), \quad \dots \quad (19)$$

which confirms the expected inverse-distance behaviour at remote points. This result is also useful in checking the validity of the far field solution to be considered in the next section.

The actual behaviour of the scattered field in the direction of the incident wave is illustrated in Figure 2, where equations (17), (18), and (19) are plotted as functions of z for the case in which $s = 10\lambda_1$. The graph shows a rather rapid

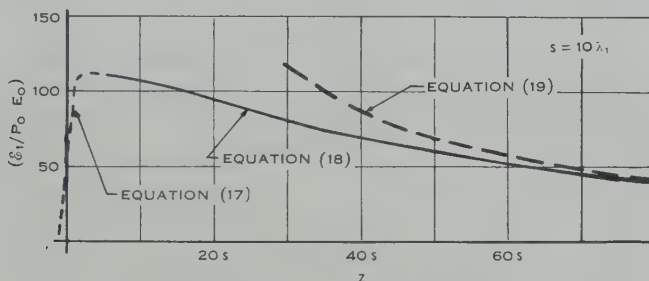


Fig. 2.—Scattered field intensity in the incident-wave direction for a 10-wavelength anomaly radius.

increase of the scattered field in the neighbourhood of the origin, to a maximum value of $|E_1| \cong 112 |p_0 E_0|$ when $z \cong 3s$, and then a gradual decrease, to become nearly identical to the inverse-distance behaviour at $z \cong 100s$. The corresponding plot for the case in which $s = 100\lambda_1$ would be similar except that the scale markings on the vertical axis would be increased by a factor of 10, and that the transition to the inverse-distance asymptote would not be adequately completed until $z \cong 1000s$. It is clear that the strongest interference between the direct and scattered fields occurs when $z \cong 3s$. The early assumption that the incident field is essentially undisturbed by the scattered field is seen to be valid, provided that $|p_0| < (\lambda_1/100s)$. This important condition is easily satisfied in all cases of tropospheric scattering of v.h.f. and microwaves.

V. SOLUTION FOR THE FAR FIELD

When the field point (x, y, z) is sufficiently far removed from the origin, the expression for ρ can be approximated as $\rho \cong r - x\xi/r - y\eta/r - z\zeta/r$ for use in the $h(\xi, \eta, \zeta)$ function in equations (8), (9), and (10). The approximation $\rho \cong r$ is adequate for use in the denominators, since each integrand is of negligibly small magnitude except near the origin. These assumptions reduce the formulas

for E_1 , E_2 , E_3 to expressions which yield to integration. The results show that, when $r \gg s$, equations (8), (9), and (10) reduce to

$$E_1 = p_0 E_0 \frac{\sqrt{\pi}}{2} \frac{s}{r} \left(\frac{2\pi s}{\lambda_1} \right)^2 \left(1 - \frac{x^2}{r^2} \right) \exp \left\{ -\frac{2\pi^2 s^2}{\lambda_1^2} \left(1 - \frac{z}{r} \right) - \frac{j2\pi}{\lambda_1} r \right\}, \quad \dots (20)$$

$$E_2 = -p_0 E_0 \frac{\sqrt{\pi}}{2} \frac{s}{r} \left(\frac{2\pi s}{\lambda_1} \right)^2 \frac{xy}{r^2} \exp \left\{ -\frac{2\pi^2 s^2}{\lambda_1^2} \left(1 - \frac{z}{r} \right) - \frac{j2\pi}{\lambda_1} r \right\}, \quad \dots (21)$$

$$E_3 = -p_0 E_0 \frac{\sqrt{\pi}}{2} \frac{s}{r} \left(\frac{2\pi s}{\lambda_1} \right)^2 \frac{xz}{r^2} \exp \left\{ -\frac{2\pi^2 s^2}{\lambda_1^2} \left(1 - \frac{z}{r} \right) - \frac{j2\pi}{\lambda_1} r \right\}. \quad \dots (22)$$

On the z -axis, these equations agree with the far field previously obtained from the exact solution. It is easy to confirm that the scattered field represented by these formulas propagates radially outward from the origin, with the electric vector perpendicular to the direction of propagation.

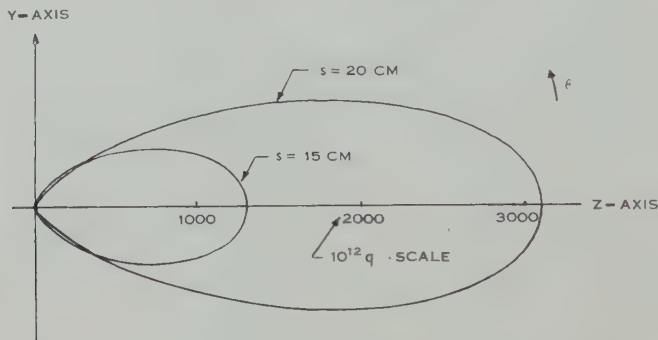


Fig. 3.—Scattered field patterns.

$$10^{12}q = 10^{12}p_0 \cdot \frac{1}{2} \sqrt{\pi} \cdot \left(\frac{s}{r} \right) \left(\frac{2\pi s}{\lambda_1} \right)^2 \exp \left\{ -\left(\frac{2\pi s}{\lambda_1} \right)^2 \sin^2 \frac{1}{2} \theta \right\},$$

$$p_0 = 10^{-6}, \lambda = 30 \text{ cm}, r = 1000 \text{ m}.$$

In order to investigate the magnitude $|\mathbf{E}| = \sqrt{(E_1^2 + E_2^2 + E_3^2)}$ of the scattered field at points far from the origin, it is convenient to introduce the polar coordinate transformation $x = r \sin \theta \cos \varphi$, $y = r \sin \theta \sin \varphi$, $z = r \cos \theta$. The ratio $q = |\mathbf{E}/E_0|$ of the scattered-field strength to the incident-field strength may then be expressed as

$$q = p_0 \frac{\sqrt{\pi}}{2} \frac{s}{r} \left(\frac{2\pi s}{\lambda_1} \right)^2 \sqrt{1 - \sin^2 \theta \cos^2 \varphi} \exp \left\{ -\left(\frac{2\pi s}{\lambda_1} \right)^2 \sin^2 \frac{\theta}{2} \right\},$$

..... (23)

which exhibits the expected inverse-distance behaviour. The angular dependence of the scattered field in the yz -plane is shown in Figure 3, where equation (23) is plotted for two different anomaly sizes.

The scattering cross section σ , being the total scattered power radiated out of a surface enclosing the scattering anomaly divided by the incident power density, is defined as

$$\sigma = \int_0^{2\pi} \int_0^\pi q^2 r^2 \sin \theta \, d\theta \, d\varphi.$$

The evaluation of this expression with the aid of equation (23) shows that the ratio $g = \sigma/\pi s^2$ of the scattering cross section σ to the "geometrical cross section" πs^2 is

$$g = \frac{\pi p_0^2}{2k^2} \{ [k^4 - k^2 + 1] - [k^4 + k^2 + 1] \exp(-2k^2) \}, \quad \dots \quad (24)$$

where $k = 2\pi s/\lambda_1$. The behaviour of this equation is shown in Figure 4, where the scaled cross-section ratio g/p_0^2 is plotted as a function of the normalized anomaly radius s/λ_1 . From equation (24) it can be shown that

$$\lim_{s \rightarrow \infty} g = (32\pi^5 p_0^2/3)(s/\lambda_1)^4,$$

which is in good agreement with the well-known Rayleigh theory of small dielectric spheres.

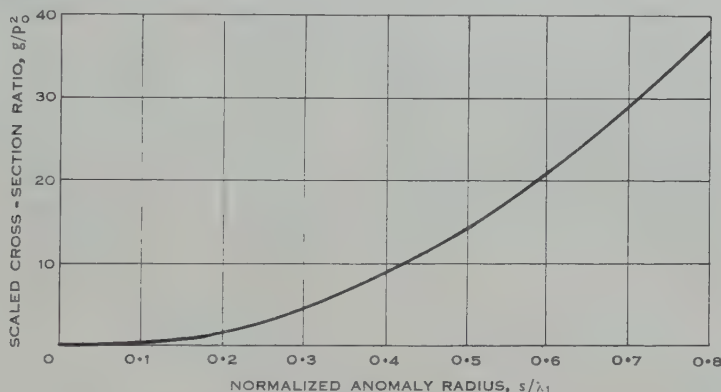


Fig. 4.—Cross-section ratio v . normalized anomaly radius.

$$g = (\pi p_0^2/2k^2) \{ (k^4 - k^2 + 1) - (k^4 + k^2 + 1) \exp(-2k^2) \},$$

where $k = 2\pi s/\lambda_1$.

VI. CONCLUSIONS

The foregoing theory describes in detail the scattering of electromagnetic energy by a single Gaussian perturbation in the refractive index. A turbulent atmosphere may be considered as being composed of many such perturbations, distributed randomly in location, size, and intensity. A statistical treatment of a large number of independent scattering elements, based on the detailed theory discussed here, should yield a worth-while contribution to the understanding of trans-horizon propagation of microwaves.

SHORT COMMUNICATIONS

LOW LATITUDE REFLECTIONS FROM THE AURORA AUSTRALIS*

By T. J. SEED† and C. D. ELLYETT†

Over the past 10 years, radio reflections from aurorae have been fairly extensively studied in the northern hemisphere (Bullough and Kaiser 1954; Booker, Gartlein, and Nichols 1955). The southern auroral zone, however, is so far removed from habitable land that very little information has yet become available on the aurora australis. Visual observations are relatively slight, and radio observations are only now commencing.

In March of this year, at the Radio Observatory of Canterbury University College, situated 14 miles from Christchurch, New Zealand, at lat. 43.6°S ., long. 172.6°E ., geographic, and lat. 47.8°S . geomagnetic, a search was started for possible radio reflections from the southern aurorae, although the station is at a very low geomagnetic latitude for such studies.

Equipment parameters used in the experiment are as follows:

Transmitter

Radio frequency	69 Mc/s
Pulse recurrence frequency	75 p.p.s.
Pulse length	20 μsec

Aerial

Polarization	horizontal
No. of half wave elements	12
Power gain over isotropic radiator	120
Azimuthal half-power points	$\pm 12^{\circ}$
Elevational half-power points	5.5 and 19.5°
Direction: great circle path to geomagnetic S. pole	195° geographic

Receiver

Detection sensitivity for P.R.F. and pulse lengths as stated	6×10^{-14} W at aerial terminals
Noise figure	4 dB absolute
Bandwidth	300 kc/s

The maximum angle of aerial elevation from which a radio reflection can be expected on Chapman's (1953) geometrical theory of auroral echoes is given by $\epsilon_{\text{max.}} = \tan^{-1} (\frac{1}{2} \tan \alpha)$, where α is the co-latitude. For the present observatory

* Manuscript received August 2, 1957.

† Canterbury University College, Christchurch, New Zealand.

$\alpha=42.2^\circ$, giving a maximum reception angle of 24° , which lies near the upper half-power point of the aerial polar diagram. Since, however, aurorae are seldom seen below a minimum height of 90 km, the maximum elevation angle may be considered to be 21° . The maximum height of echo reflection will be about 250 km, corresponding to an angle ϵ which has fallen to the lower half-power point.

A graph showing the range-height relationship in terms of aerial elevation angle is given in Figure 1. Following Chapman (1953), it is assumed that any echo region will lie along the Earth's magnetic lines of force, and will be perpendi-

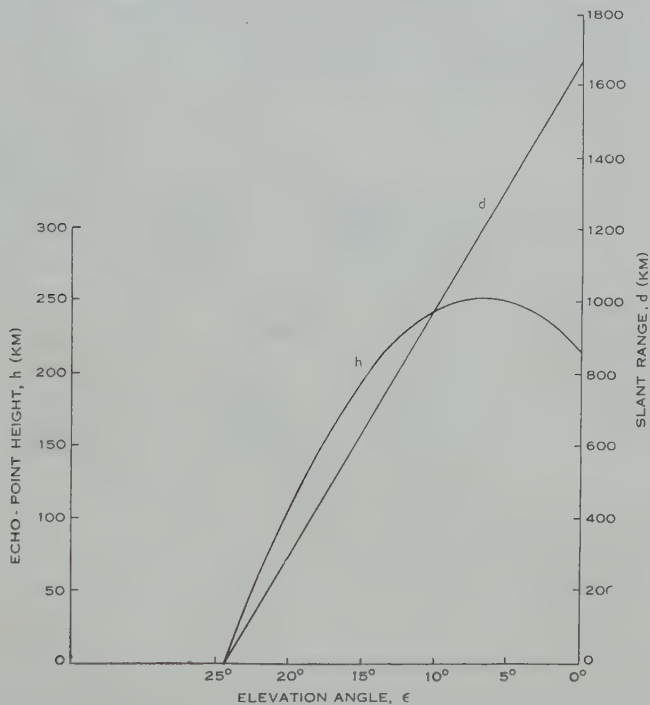


Fig. 1.—Auroral echo heights and slant ranges versus aerial elevation angle, computed for geomagnetic latitude 47.8°S .

cular to the radio beam. In this case long range will be associated with great height of the echo region, and a low pulse recurrence frequency of 75 p.p.s. was chosen to allow a time base range of approximately 2000 km, in case echoes were received from aurorae up to F -region altitudes in the ionosphere. No persistent F -region auroral echoes have yet been detected at the observatory.

Auroral radio echoes were recorded continuously at Christchurch on March 10, 1957 between 1341 and 1626 hr U.T. at slant ranges of 325–500 km, despite heavy interference. Heights of 116–165 km corresponding to these ranges are inferred from the Chapman theory. This aurora was observed visually (in generally overcast conditions) over the southern part of New Zealand, Campbell Island, and Australia between 1030 and 1130 U.T., and over the full length of New Zealand between 1300 and 1400 U.T. The display was again visible in

Australia between 1600 and 1630 U.T. (Thomsen, personal communication 1957).

The next radar display was observed on April 10, commencing at 0659 U.T. just after the equipment had been switched on. Comparison with collated visual records (Thomsen, personal communication 1957) and magnetic records (Cullington, personal communication 1957) is given in pictorial representation in Figure 2.

This shows a clear correspondence between the visual and the radar observations. In general an exact correspondence cannot be expected, since the great circle path from the aerial looks over an unbroken 3000 km of the South Pacific Ocean to the geomagnetic pole and visual observers are on the average some 100 km to the west of this path.

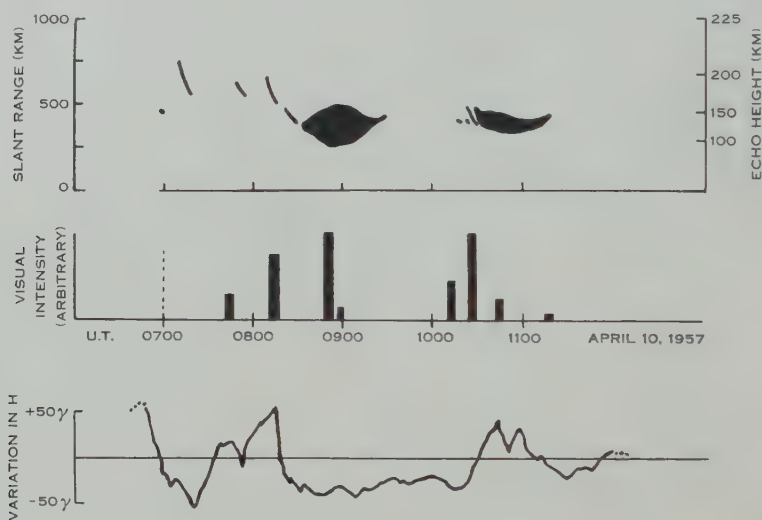


Fig. 2.—Radar, visual, and magnetic correlation for the aurora of April 10, 1957.

The shortest radar range recorded during this display was 250 km, corresponding to a height of 93 km. (The range decreased at the commencement and increased at the conclusion of each of the two main epochs.) Comparison with a record of the *D*-component of the Earth's magnetic field taken on a high speed flux-gate magnetometer some 40 km north of the radar station (Cullington, personal communication) on this occasion supported the observations of Meek (1954) that the commencement of an auroral return coincides with a marked gradient in the magnetogram.

A number of subsequent auroral echoes have now been obtained, as summarized in Table 1.

It is clear from Table 1 that both major and minor displays of the aurora australis can be recorded by radar methods from the low geomagnetic latitude of 47.8°S . Aurorae Nos. 3–10 in Table 1, in general smaller than No. 2 (of April 10, 1957) have not shown any obvious correlation with magnetograms.

The occurrence of 10 aurorae in 61 days' observing is much higher than was anticipated for this latitude from a survey of northern hemisphere V.H.F. observations.

TABLE 1

SUMMARY OF V.H.F. AURORAL RADAR OBSERVATIONS MARCH 5 TO JUNE 11, 1957

1957 Observing Periods		Radiated Pulse Power (kW)	Date of Aurora	Periods of Detection (U.T.)	Min. Range and Extent (km)	Computed Echo Height (ex Chapman) (km)
Date	Time (hr U.T.)					
Mar. 5-11	0745-1800	70	Mar. 10	(described in text)		
Apr. 3-12	0700-1800	70	Apr. 10	Fig. 2		
Apr. 13-27	0600-1800					
May 1,3,4	0600-1800					
May 6-8	0600-1800	150	May 7	0920-0921	380	134
				0924-0938	340	120
				0940-0943	320	117
May 12-31	0600-1800	150	May 12	1526-1532	370	131
		150	May 13	1630-1645	340-290	120-105
		150	May 20	0805-0813	275	100
				1441-1447	350	125
		150	May 31	0925-0935	440	152
				1530-1733	350-420	125-146
June 1-4	0530-1800	150	June 2	1332-1354	360-380	128-134
		150	June 4	1035-1047	350-400	125-140
June 6,10,11	0530-1800	150	June 10	0535-0551	300-330	110-120
				0558-0625	330	120
				0635-0646	In patches at 350 km	125

The authors are indebted to the New Zealand I.G.Y. Committee for financial assistance with the project. They also wish to thank Mr. I. L. Thomsen, Director of the Carter Observatory, Wellington, for supplying the summary of visual auroral observations, and Mr. A. L. Cullington, Director of the Christchurch Magnetic Observatory, for making available magnetogram results.

References

- BOOKER, H. G., GARTLEIN, C. W., and NICHOLS, B. (1955).—*J. Geophys. Res.* **60**: 1.
 BULLOUGH, K., and KAISER, T. R. (1954).—*J. Atmos. Terr. Phys.* **5**: 189.
 BULLOUGH, K., and KAISER, T. R. (1955).—*J. Atmos. Terr. Phys.* **6**: 198.
 CHAPMAN, S. (1953).—*J. Atmos. Terr. Phys.* **3**: 1.
 MEEK, J. H. (1954).—*J. Geophys. Res.* **58**: 328.

THE DISTRIBUTION OF FLARE HEIGHTS AS DERIVED FROM LIMB FLARES*

By R. G. GIOVANELLI† and MARIE K. MCCABE†

In recent papers, J. W. Warwick (1955) and Constance S. Warwick (1955) have discussed the heights of flares—the former by examining the distribution of apparent areas across the disk, and the latter by measuring heights of limb flares. Both analyses indicate that flares have heights of the order of $10\text{--}20 \times 10^3$ km; J. W. Warwick's height-frequency distribution is well represented by $\exp(-\beta h)$, where $\beta = 0.492 \times 10^{-5}$ km, while C. S. Warwick's height distribution has a maximum at about 14×10^3 km.

While limb observations would appear to provide the more attractive method for deriving the flare height distribution, one of C. S. Warwick's difficulties has been the small number, 39, of limb flare observations available. Observations with the Sydney Lyot monochromator from April 6, 1956 to April 5, 1957 have enabled us to detect 90 flares whose tops projected beyond the limb, and this larger sample has been used for a flare height analysis in which uncertainties in the position of the base in front of or behind the limb have been avoided. Further, we have been able to show that the limb flares are the same type of event as observed on the disk, and this has enabled us to find the mean flare area corresponding to the derived mean height. Our observations include smaller flares than those analysed by Miss Warwick, so we may expect our mean height also to be less.

Observational Material

The Sydney Lyot monochromator is operated on a $\frac{1}{2}$ min cycle, a 16 mm diameter H α image of the disk being recorded on Eastman Kodak IV-E film. While the small image size militates against high resolution, these observations allow all flares of area above about 20×10^{-6} of the Sun's hemisphere to be detected; most of the flares are, in fact, very small, of class 1-.

There are two main difficulties associated with the analysis of heights of flares at the limb:

(a) the identification of events at the limb as flares similar to those which are observed on the disk. In the present analysis we shall demonstrate the validity of the identification by comparing rates of limb and disk occurrences of flares;

(b) the measurement of the true height of a flare when the base of the flare is at an unknown or uncertain distance behind or in front of the limb. It is possible to avoid this difficulty by measuring only the distance by which a flare

* Manuscript received October 11, 1957.

† Division of Physics, C.S.I.R.O., University Grounds, Chippendale, N.S.W.

projects beyond the limb, irrespective of the position of its base, and using an appropriate analysis as described below.

During the above-mentioned 12 months' interval, covering a period around sunspot maximum, 90 events, classified as flares on the basis of their brightnesses and lifetimes, were observed with the top of the flare region projecting beyond the limb. For present purposes, the top of the flare has been taken as the limit of the region considered to be of flare brightness. We have specifically excluded fainter parts of brightness equal to those of the chromosphere or of the brighter projections such as stable plages, as well as the bright diffuse regions, ejected by some flares, which eventually fade to become dark surges. As a consequence, we believe that this analysis provides the height distribution for what are commonly accepted as normal flares.

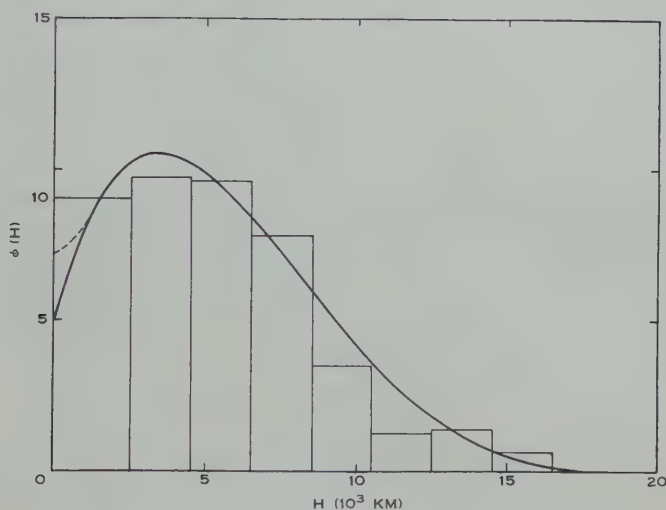


Fig. 1.—Histogram showing $\varphi(H)$, the observed number of flares, per 10^3 km projecting height H above the limb. The continuous curve shows the analytic function used to represent $\varphi(H)$, the dotted portion between $0 \leq H \leq 1.19 \times 10^3$ km being the correction required to the analytic function to result in a zero value for the derived function $f(h)$ over the same range $0 \leq h \leq 1.19 \times 10^3$ km.

The observed height distribution of the parts of these flares projecting beyond the chromosphere is shown in the histogram of Figure 1; the uncertainty of the individual measurements is believed to be of the order of 2×10^3 km.

Reduction of Observations

Let there be N flares per radian of longitude, of which $Nf(h)dh$ have tops in a true height range h to $h+dh$. If the radius vector from the flare to the centre of the Sun makes an angle θ with the perpendicular to the line of sight, the flare overlaps the limb by a projected distance

$$H = (R + h) \cos \theta - R, \quad \dots \dots \dots (1)$$

where R is the solar radius (Fig. 2).

Assuming that the Earth and the flares lie in the Sun's equatorial plane (a simplification that introduces little error), the number of flares in $d\theta$, dh is given by $Nf(h)dh d\theta$, corresponding to which there will be $Nf(h)dH \sec \theta d\theta$ projecting beyond the limb in a projected height range dH . The total number in dH is thus

$$\varphi(H)dH = 2NdH \int_0^{\pi/2} f(h) \sec \theta d\theta. \quad \dots\dots\dots (2)$$

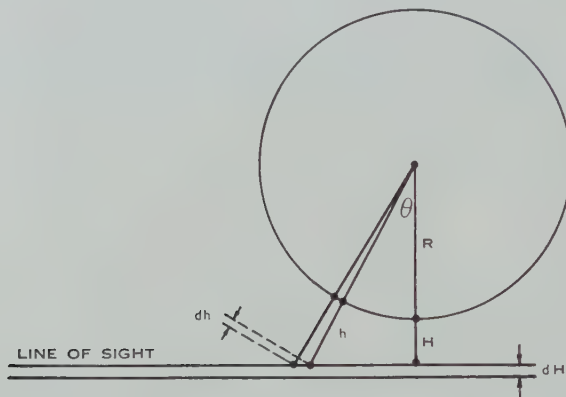


Fig. 2.—The geometrical relationship between the observed projected height H and the true height of a flare, h .

But the line of sight, (1), gives on differentiation

$$dh = (R + h) \tan \theta d\theta,$$

whence (2) becomes

$$\begin{aligned} \varphi(H) &= 2N \int_H^\infty \frac{f(h)dh}{(R+h) \sin \theta} \\ &= 2N \int_H^\infty \frac{f(h)dh}{\{(2R+h+H)(h-H)\}^{\frac{1}{2}}}. \end{aligned}$$

Since $h+H \ll 2R$ for all significant cases,

$$\varphi(H) = \frac{2N}{(2R)^{\frac{1}{2}}} \int_H^\infty \frac{f(h)dh}{(h-H)^{\frac{1}{2}}}. \quad \dots\dots\dots (3)$$

The solution of this integral equation for the flare height distribution $f(h)$ in terms of the radius R , the observed number of limb flares N , and the observed height distribution of the parts of flares projecting beyond the limb, $\varphi(H)$ (see Fig. 1), is obtained by assuming $f(h)$ expressible, over the range $0 \leq h \leq K$ in which $f(h)$ has significant values, by a polynomial

$$f(h) = \sum a_n h^n,$$

and that $f(h) = 0$, $h > K$. With a suitable change of variable, (3) then becomes

$$\varphi(H) = \alpha \int_0^{K-H} \sum a_n (x+H)^n \cdot x^{-\frac{1}{2}} dx, \quad \dots\dots\dots (4)$$

where $\alpha = N(2/R)^{\frac{1}{2}}$.

The integral in (4) can be evaluated term by term, yielding a series of half-integral powers of $(K-H)$. If we fit $\varphi(H)/(K-H)^{\frac{1}{2}}$ by a power series $\sum b_n H^n$, we can equate the coefficients of equal powers of H on either side of the expression resulting from (4), and these lead to a solution for the a_n in terms of the known b_n .

Now, the observed function $\varphi(H)$ is such that it drops effectively to zero at a height $H \approx 17.5 \times 10^3$ km, so we choose $K = 17.5 \times 10^3$ km. The function $\varphi(H)/(K-H)^{\frac{1}{2}}$ can be fitted by a polynomial of the fourth degree in H to within the limits of observational error, the corresponding curve $(K-H)^{\frac{1}{2}} \sum b_n H^n$ being superimposed on the histogram of Figure 1.

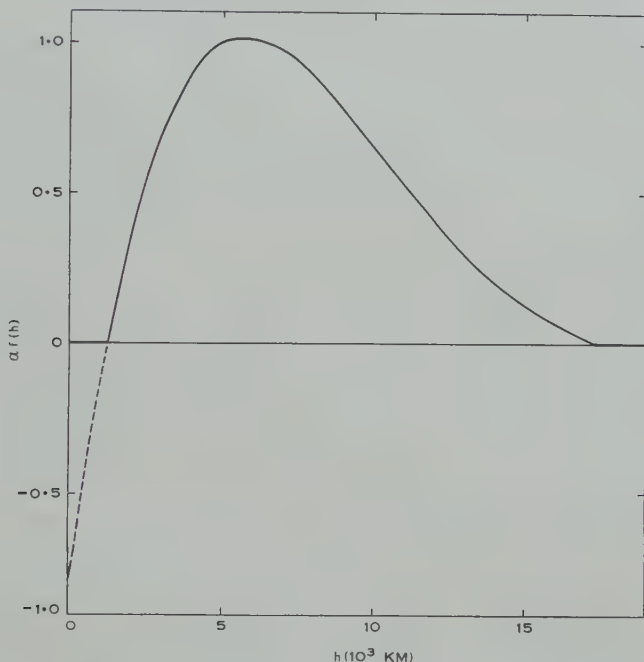


Fig. 3.—The derived flare height distribution $f(h)$ as a function of h . The ordinates are values of $\alpha f(h)$, where $\alpha = N(2/R)^{\frac{1}{2}}$.

The resulting solution, $f(h) = \sum a_n h^n$, is given in Figure 3. We may note that this function has negative values (dotted curve) near the origin, indicating the inadequacy of the power series $\sum b_n H^n$ near the origin. Actually, this is of little significance; by allocating zero values to $f(h)$ for heights below 1.19×10^3 km we affect only the correspondence with observed heights in the range $0-1.19 \times 10^3$ km, and in fact the required observed value of $\varphi(H)$ at $H=0$ is 7.2 , in better agreement with the histogram than the analytical curve at this point.

Discussion of Results

This analysis of limb flares leads to the real height distribution as shown in Figure 3, with flare heights lying between 1.2×10^3 and 17.5×10^3 km. The mean flare height is 7.3×10^3 km; the extremities of the distribution are unreliable because of the lack of resolving power and the limited observational data.

From these results, we can now calculate N , the number of flares per radian of longitude, with the result $N=331$. The total number of observable flares occurring in the period of observation should, on this basis, be πN + one-half the number of limb flares (those with bases behind the limb). The expected total is 1084, with a standard deviation, derived from a total of 90 limb flares, of $\sigma=\pm 115$; the analysis and the non-random nature of flare occurrences—in that they are associated with active centres—would increase σ somewhat. This total may be compared with an actual total of 1327 flares observed during the

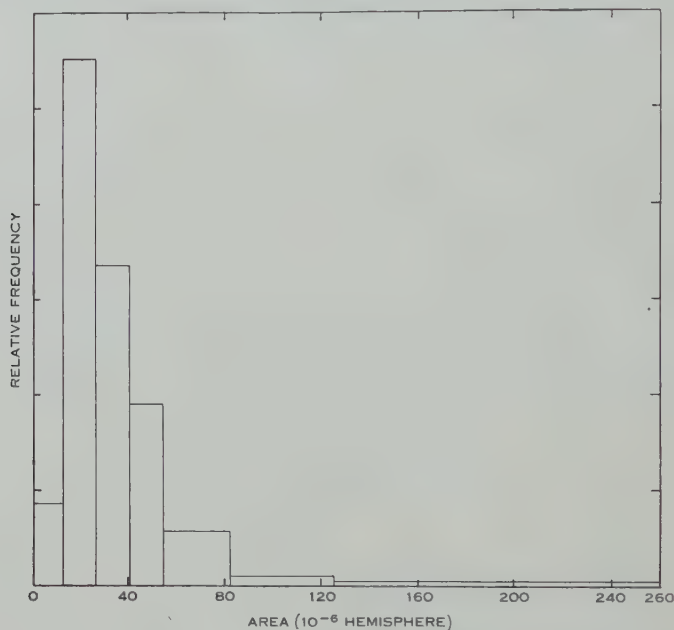


Fig. 4.—Histogram showing apparent area distribution for 269 flares occurring within 30° of the centre of the Sun's disk.

same period. The agreement is sufficiently good (within about 2σ) to justify our conclusion that the flares observed on the limb are the same type of event as those observed on the disk. In this case the area distribution applicable to flares near the centre of the disk should also be applicable to our limb flares. Figure 4 shows a histogram giving our observed distribution of apparent area for 269 flares occurring within 30° of the centre of the Sun's disk during part of the observation period; from this we deduce that the average area of these flares is about 40×10^{-6} of the Sun's hemisphere ($1.2 \times 10^8 \text{ km}^2$). As shown above, such a flare has an average height of $7.3 \times 10^3 \text{ km}$.

References

- WARWICK, CONSTANCE S. (1955).—*Astrophys. J.* **121**: 385.
 WARWICK, J. W. (1955).—*Astrophys. J.* **121**: 376.

DECAY TIME OF THE LUMINESCENCE OF A ZINC SULPHIDE NEUTRON DETECTOR FOR NEUTRON AND γ -RAY EXCITATION*

By G. M. BAILEY† and J. R. PRESCOTT‡

In experiments being conducted in this laboratory on $(d,n\gamma)$ reactions using fast coincidence techniques, an attempt was made to use phosphors containing zinc sulphide as the neutron detector.

When it was found that this type of detector did not produce the discrimination expected between neutrons and γ -rays, its response to both types of radiation was investigated in detail.

Experimental Technique

The neutron detector consisted of $\text{ZnS}(\text{Ag})$ § moulded in a Lucite cylinder 1 in. in diameter by $\frac{3}{4}$ in. thick (Hornyak 1952), mounted directly on the face of a photomultiplier tube. Dumont 6292 and R.C.A. 6342 tubes were used, the first chosen for its low noise level, the second for its faster response.

A direct decay measurement for γ -excitation using a small photomultiplier collector time constant was not possible owing to the small pulse size obtained and the presence of a large background due to the long-term phosphorescence of the phosphor. Instead, the decay time was determined from a measure of the integrated pulse rise time.

The collector current pulses from individual scintillations were integrated by a long time constant ~ 10 μsec , and fed from a cathode follower into two cascaded wide band amplifiers (total rise time 3.6 μsec), and finally into the vertical amplifier of a Tektronix 545 oscilloscope (rise time 12 μsec). Pulse rise times were determined photographically from a short time exposure spectrum. The oscilloscope sweep was later calibrated using a 10 Mc/s oscillator.

The form of the output pulse appearing across the load R , for an exponentially decaying scintillator having a time constant T , when $RC \gg T$, is

$$V_0(t) = (i_0 T / C) [1 - \exp(-t/T)], \quad \dots\dots\dots (1)$$

C being the stray capacitance of the collector circuit to ground. The rise time (10–90 per cent.) of such a pulse is $2.2T$. Allowance must be made for the

* Manuscript received November 4, 1957.

† Physics Department, University of Melbourne.

‡ Australian Atomic Energy Commission, University of Melbourne; present address: University of British Columbia, Vancouver, B.C., Canada.

§ Type G-86, Levy-West Laboratories, Wembley.

apparatus rise time T_a , and, if T_m is the measured rise time, it can be assumed that

$$T_m^2 = T_a^2 + K^2 T^2, \dots\dots\dots (2)$$

K being a constant approximately equal to the theoretical value 2.2. If relation (2) is valid, a plot of $\log (T_m^2 - T_a^2)$ against $\log T^2$ for different exponentially decaying scintillators should be a straight line of slope 1, and intercept $\log K^2$ on the $\log (T_m^2 - T_a^2)$ axis.

The validity of this expression was tested by measuring the rise times of several organic scintillators which have known decay constants. The results are shown plotted in Figure 1; the points lie closely on a straight line of slope 1 as predicted. The value of K is found to be 2.1.

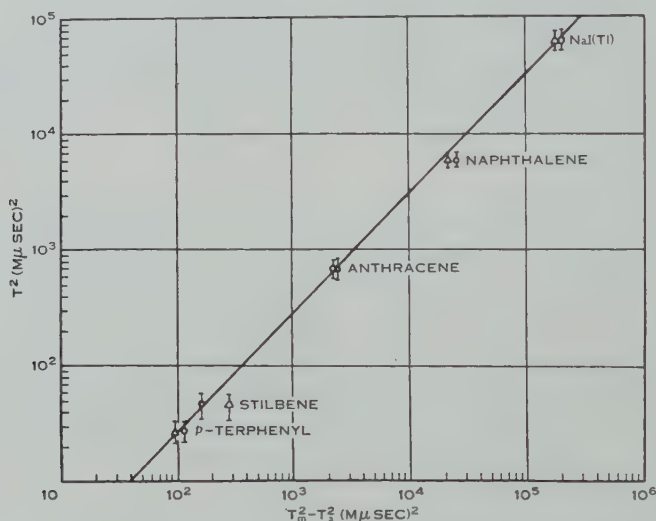


Fig. 1.—Square of scintillator decay time plotted as a function of square of effective rise time, for a number of scintillators under γ -ray excitation. \triangle Dumont 6292; \circ R.C.A. 6342.

In choosing the value of the decay constants for the various scintillators it was realized that the physical dimensions could be of importance (Birks and Little 1953). This factor largely determines the variance of values quoted in the literature. A large variance was evident in the case of naphthalene and anthracene and so these were measured using the technique of Bittman, Furst, and Kallmann (1952). In the other cases the mean of values quoted by Liebson, Bishop, and Elliot (1950), Lundby (1950), Bittman, Furst, and Kallmann (1952), and Swank and Buck (1955) was used.

Decay of ZnS(Ag)+Lucite for γ -Ray Excitation

Measurements of pulses from the detector when excited by γ -rays from ^{60}Co indicated a fast decay. The resulting decay constants for the two tubes used, taken from the calibration line, were Dumont 6292: $8.4 \pm 2.3 \mu\text{sec}$,

R.C.A. 6342 : 8.5 ± 2.7 μ sec, giving a mean of 8.5 ± 2.5 μ sec. The accuracy of the measurements of rise times was limited by the fluctuating nature of the pulses, due partly to the long-term phosphorescence of the ZnS(Ag) and partly to the superimposed photomultiplier noise ; this latter effect being more noticeable with the 6342 tube.

Decay of the Scintillator for 2.6 MeV Neutrons

Using neutrons produced by the $^2\text{H}(^2\text{H},n)^3\text{He}$ reaction from the University of Melbourne 750 keV electrostatic generator and the technique described above, it was found that the pulse rise time from the zinc sulphide was very much slower than that observed for γ -rays. On account of the relatively slow decay, a direct measurement of the decay form was possible.

The decay appeared complex, but from a number of photographs of the decay it was found to be satisfactorily represented by a combination of two exponential decay regions, together with a very much longer decaying tail of a few microseconds. No initial fast decay component was observed, and the results gave decay constants : 0.13 ± 0.02 , 0.34 ± 0.04 μ sec for 2.6 MeV neutrons.

On closer analysis a hyperbolic decay for the light intensity of the form

$$I(t) = I_0 / (1 + kt)^\beta \quad \dots\dots\dots (3)$$

could be fitted to the results for the first microsecond over which it was estimated 75 per cent. of the total light intensity was emitted. The exponent β was found to be closely one.

Conclusions

The decay time of a ZnS(Ag) + Lucite neutron detector has been reported to have a principal decay constant of 40 μ sec for fast neutrons, whilst ZnS(Ag) in powder form has this same decay time for γ - and α -particle excitation (Koontz 1953 ; Koontz, Keepin, and Ashley 1955).

The results obtained for neutron excitation, whilst being in disagreement with Koontz, agree with the decay form observed by several investigators for α -particle excitation of ZnS(Ag) (Bittman, Furst, and Kallmann 1952 ; Hornyak 1952 ; Emmerich 1954 ; Smidt 1955). This is to be expected since the detection process takes place mainly from (n,p) scattering in the Lucite, and the $^{32}\text{S}(n,p)^{32}\text{P}$ reaction in the phosphor.

The rapid decay for γ -bombardment has been noted by Emmerich (1954), and points to the absence of the slow recombination process between diffusing positive holes and electrons which takes place when the phosphor is excited by highly ionizing particles (Smidt 1955).

The above work was carried out as part of a joint programme of the Physics Department, University of Melbourne, and the Australian Atomic Energy Commission. Thanks are due to Professor Sir Leslie Martin for affording us the facilities of the laboratory. One of us (G.M.B.) was assisted by a research grant from the University of Melbourne.

References

- BIRKS, J. B., and LITTLE, W. A. (1953).—*Proc. Phys. Soc. Lond. A* **66**: 921.
 BITTMAN, L., FURST, M., and KALLMANN, H. (1952).—*Phys. Rev.* **87**: 83.
 EMMERICH, W. S. (1954).—*Rev. Sci. Instrum.* **25**: 69.
 HORNYAK, W. F. (1952).—*Rev. Sci. Instrum.* **23**: 264.
 KOONTZ, P. G. (1953).—A.E.C.U. Rep. 2913.
 KOONTZ, P. G., KEEPIN, G. R., and ASHLEY, J. E. (1955).—*Rev. Sci. Instrum.* **26**: 352.
 LIEBSON, S. H., BISHOP, M. E., and ELLIOT, J. O. (1950).—*Phys. Rev.* **80**: 907.
 LUNDBY, A. (1950).—*Phys. Rev.* **80**: 477.
 SMIDT, V. D. (1955).—*Ann. Phys., Lpz.* **15**: 325.
 SWANK, R., and BUCK, W. (1955).—*Rev. Sci. Instrum.* **26**: 15.

COULOMB WAVE FUNCTIONS*

By A. LEARNER† and B. A. ROBSON‡

Introduction

Coulomb wave functions are required for the solution of many physical problems involving charged particles. Recently, Froberg's (1955) review article has shown the need for a skeleton table of the values corresponding to the higher incident energies now so readily available. The region of interest is that formed by the parallelogram with corners in $(\rho, \eta) = (10, 0), (20, 5), (20, 15),$ and $(10, 10)$. In the present paper, such a table of values for f_L and f'_L (defined in equations (3c) and (4a)) is given for $\eta \leq 10, L = 5, 6, 10,$ and $11, \Delta\rho = 1$ and $\Delta\eta = 1$. These integrals were calculated on the CSIRAC (Physics Department, University of Melbourne). Work is in progress to obtain the corresponding values for g_L and g'_L (defined in equations (3d) and (4b)).

Notation

The differential equation is

$$d^2y/d\rho^2 + \{1 - 2\eta/\rho - L(L+1)/\rho^2\}y = 0, \quad \dots\dots\dots (1)$$

which has a regular solution $F_L(\eta, \rho)$ and a logarithmic solution $G_L(\eta, \rho)$, which are normalized to act as sine-cosine functions as $\rho \rightarrow \infty$,

$$F_L(\eta, \rho) \sim \sin \theta_L, \quad \dots\dots\dots (2a)$$

$$G_L(\eta, \rho) \sim \cos \theta_L, \quad \dots\dots\dots (2b)$$

where

$$\theta_L = \{\rho - \eta \ln 2\rho - L\pi/2 + \sigma_L\}, \quad \dots\dots\dots (2c)$$

$$\sigma_L = \arg \Gamma(i\eta + L + 1). \quad \dots\dots\dots (2d)$$

* Manuscript received December 5, 1957.

† Mathematics Department, University of Melbourne; present address: Trinity College, Cambridge.

‡ Physics Department, University of Melbourne.

These functions are given by the integral representation (Bloch *et al.* 1950)

$$F_L = A_L \rho^{L+1} f_L, \quad \dots\dots\dots (3a)$$

$$G_L = A_L \rho^{L+1} g_L, \quad \dots\dots\dots (3b)$$

where

$$f_L = \int_0^1 (1-z^2)^L \cos (2\eta \operatorname{artanh} z - \rho z) dz, \quad \dots\dots\dots (3c)$$

$$g_L = \int_0^1 (1-z^2)^L \sin (2\eta \operatorname{artanh} z - \rho z) dz \\ + e^{\pi\eta} \int_0^\infty (1+u^2)^L \exp (-u\rho - 2\eta \operatorname{artan} 1/u) du, \quad \dots\dots (3d)$$

and

$$A_L = 2(L^2 + \eta^2)^{-\frac{1}{2}} A_{L-1}, \quad \dots\dots\dots (3e)$$

$$A_0 = (1 - e^{-2\pi\eta})^{\frac{1}{2}} (2\pi\eta)^{-\frac{1}{2}} = \{C_0(\eta)\}^{-1}. \quad \dots\dots\dots (3f)$$

The derivatives f'_L ($\partial F_L / \partial \rho$) and g'_L are thus

$$f'_L = \int_0^1 z(1-z^2)^L \sin (2\eta \operatorname{artanh} z - \rho z) dz, \quad \dots\dots\dots (4a)$$

$$g'_L = - \int_0^1 z(1-z^2)^L \cos (2\eta \operatorname{artanh} z - \rho z) dz \\ - e^{\pi\eta} \int_0^\infty u(1+u^2)^L \exp (-u\rho - 2\eta \operatorname{artan} 1/u) du. \quad \dots\dots (4b)$$

Use and Accuracy of Tables

The required values of F_L and F'_L are obtained by using (3a), (3e), and (3f) with the tabulated f_L , f'_L and noting that

$$F'_L = A_L \rho^{L+1} f'_L + (L+1) F_L / \rho. \quad \dots\dots\dots (5)$$

$C_0(\eta)$ may be obtained from the National Bureau of Standards tables (1952). For convenience, the tabulated values of L are adjacent 5, 6 and 10, 11 so that intermediate and other L values may be obtained by the recurrence formulae

$$\rho^2 L f_{L+1} = 2(2L+1) \{ \eta \rho + L(L+1) \} f_L - 4(L+1) (L^2 + \eta^2) f_{L-1}, \quad \dots (6a)$$

and

$$\rho L f'_L = 2(L^2 + \eta^2) f_{L-1} - \{ \eta \rho + L(2L+1) \} f_L. \quad \dots\dots\dots (6b)$$

These formulae permit a check on the accuracy of the tables, which was generally found to be at least 0.01 per cent. For the transition line $\rho = 2\eta$, $2\eta \operatorname{artanh} Z - \rho Z \simeq 0$ and one cannot expect a sensible answer for the method of computation used. To the left of the transition line F_L oscillates rapidly and this causes round-off errors to be important. Any error in F_L means an admixture of G_L , and since G_L becomes large as ρ decreases, the corresponding error in F'_L increases rapidly. For $\eta = 10$, $\rho = 10$, $L = 5$ and 6, the possible error is ~ 0.1 per cent. This represents the maximum error in the table except near a zero of f_L . Similar comments apply to f'_L .

TABLE 1
VALUES OF f_L AND f'_L

ρ	$L=5$		$L=6$		$L=10$		$L=11$	
	f_5	f'_5	f_6	f'_6	f_{10}	f'_{10}	f_{11}	f'_{11}
$\eta=10$								
10	0.01157	0.006781	0.01492	0.008046	0.02973	0.011845	0.03334	0.012424
11	0.02024	0.010808	0.02497	0.012242	0.04347	0.015715	0.04756	0.016104
12	0.03365	0.016251	0.03981	0.017629	0.06127	0.019930	0.06555	0.019918
13	0.05320	0.023042	0.06056	0.024009	0.08332	0.024122	0.08734	0.023595
14	0.08005	0.030763	0.08799	0.030866	0.10934	0.027779	0.11253	0.026636
15	0.11475	0.038547	0.12216	0.037323	0.13850	0.030290	0.14022	0.028510
16	0.15673	0.045078	0.16210	0.042185	0.16933	0.031030	0.16899	0.028698
17	0.20394	0.048710	0.20554	0.044082	0.19978	0.029454	0.19691	0.026781
18	0.25260	0.047743	0.24884	0.041734	0.22734	0.025232	0.22176	0.022538
19	0.29743	0.040836	0.28729	0.034272	0.24935	0.018349	0.24122	0.016026
$\eta=9$								
10	0.03254	0.016329	0.03895	0.017819	0.06121	0.020245	0.06561	0.020220
11	0.05231	0.023429	0.06003	0.024474	0.08363	0.024541	0.08774	0.023972
12	0.07974	0.031532	0.08807	0.031631	0.11010	0.028261	0.11332	0.027046
13	0.11540	0.039691	0.12314	0.038335	0.13974	0.030763	0.14142	0.028889
14	0.15866	0.046462	0.16416	0.043295	0.17100	0.031400	0.17051	0.028969
15	0.20726	0.050064	0.20866	0.045050	0.20173	0.029626	0.19862	0.026872
16	0.25711	0.048688	0.25274	0.042266	0.22934	0.025127	0.22346	0.022395
17	0.30251	0.040972	0.29138	0.034100	0.25108	0.017931	0.24264	0.015631
18
19	0.35354	0.006207	0.33102	0.002682	0.26752	—0.002434	0.25626	—0.002753
20	0.34804	—0.017509	0.32375	—0.017371	0.25946	—0.013644	0.24853	—0.012672
$\eta=8$								
10	0.07940	0.032379	0.08817	0.032468	0.11092	0.028766	0.11417	0.027471
11	0.11613	0.040952	0.12422	0.039436	0.14106	0.031250	0.14267	0.029274
12	0.16079	0.047974	0.16641	0.044483	0.17276	0.031768	0.17210	0.029235
13	0.21091	0.051505	0.21203	0.046050	0.20375	0.029775	0.20038	0.026942
14	0.26199	0.049627	0.25688	0.042755	0.23137	0.024980	0.22516	0.022218
15	0.30788	0.040972	0.29563	0.033798	0.25280	0.017459	0.24405	0.015194
16
17	0.35646	0.003756	0.33284	0.000623	0.26776	—0.003380	0.25638	—0.003541
18	0.34804	—0.020791	0.32323	—0.019904	0.25871	—0.014619	0.24784	—0.013466
19	0.31517	—0.044420	0.29346	—0.039144	0.23891	—0.024714	0.22980	—0.022361
20	0.26095	—0.062834	0.24646	—0.053894	0.21007	—0.032485	0.20378	—0.029264
$\eta=7$								
10	0.16318	0.049630	0.16888	0.045753	0.17460	0.032129	0.17375	0.029490
11	0.21492	0.053032	0.21567	0.047072	0.20584	0.029896	0.20219	0.026987
12	0.26726	0.050532	0.26127	0.043175	0.23343	0.024786	0.22688	0.022002
13	0.31355	0.040784	0.30002	0.033330	0.25450	0.016930	0.24542	0.014716
14
15	0.35907	0.000945	0.33438	—0.001662	0.26786	—0.004360	0.25641	—0.004349
16	0.34739	—0.024348	0.32224	—0.022577	0.25782	—0.015596	0.24705	—0.014257
17	0.31095	—0.047857	0.28986	—0.041574	0.23712	—0.025507	0.22828	—0.022997
18	0.25377	—0.065174	0.24082	—0.055436	0.20765	—0.032939	0.20175	—0.029629
19	0.18385	—0.072927	0.18152	—0.061821	0.17234	—0.037105	0.16995	—0.033514
20	0.11159	—0.069793	0.11999	—0.059890	0.13465	—0.037727	0.13575	—0.034411

TABLE 1 (Continued)

ρ	$L=5$		$L=6$		$L=10$		$L=11$	
	f_5	f'_5	f_6	f'_6	f_{10}	f'_{10}	f_{11}	f'_{11}
$\eta=6$								
10	0.27298	0.051364	0.26594	0.043495	0.23550	0.024539	0.22861	0.021746
11	0.31949	0.040349	0.30453	0.032659	0.25615	0.016341	0.24675	0.014193
12
13	0.36122	-0.002259	0.33552	-0.004177	0.26783	-0.005367	0.25631	-0.005174
14	0.34598	-0.028153	0.32069	-0.025357	0.25679	-0.016567	0.24614	-0.015040
15	0.30585	-0.051265	0.28569	-0.043929	0.23521	-0.026256	0.22667	-0.023599
16	0.24589	-0.067162	0.23478	-0.056716	0.20516	-0.033323	0.19968	-0.029944
17	0.17499	-0.072814	0.17492	-0.061635	0.16969	-0.037055	0.16774	-0.033491
18	0.10396	-0.067513	0.11432	-0.058261	0.13225	-0.037266	0.13372	-0.034070
19	0.043023	-0.053141	0.060734	-0.047946	0.096220	-0.034323	0.10051	-0.031937
20	-0.000601	-0.033656	0.019830	-0.033446	0.064397	-0.029015	0.070548	-0.027705
$\eta=5$								
11	0.36278	-0.005870	0.33621	-0.006914	0.26765	-0.006395	0.25610	-0.006011
12	0.34366	-0.032146	0.31852	-0.028199	0.25560	-0.017520	0.24513	-0.015808
13	0.29980	-0.054519	0.28093	-0.046134	0.23318	-0.026953	0.22498	-0.024161
14	0.23740	-0.068662	0.22842	-0.057663	0.20263	-0.033632	0.19758	-0.030204
15	0.16613	-0.072029	0.16837	-0.061040	0.16707	-0.036929	0.16554	-0.033412
16	0.09697	-0.064631	0.10908	-0.056292	0.12996	-0.036754	0.13176	-0.033686
17	0.039623	-0.049051	0.057932	-0.045186	0.094584	-0.033549	0.099049	-0.031344
18	0.000162	-0.029644	0.019850	-0.030699	0.063594	-0.028156	0.069735	-0.027026
19	-0.019993	-0.011204	-0.003454	-0.016166	0.038653	-0.021631	0.045380	-0.021568
20	-0.023861	0.002448	-0.013409	-0.004381	0.020357	-0.015026	0.026696	-0.015803
$\eta=4$								
10	0.34029	-0.036223	0.31570	-0.031033	0.25427	-0.018447	0.24400	-0.016554
11	0.29283	-0.057463	0.27564	-0.048105	0.23106	-0.027590	0.22322	-0.024679
12	0.22847	-0.069532	0.22184	-0.058215	0.20007	-0.033861	0.19546	-0.030407
13	0.15752	-0.070520	0.16203	-0.060026	0.16450	-0.036725	0.16338	-0.033279
14	0.09090	-0.061225	0.10442	-0.054037	0.12778	-0.036187	0.12988	-0.033265
15	0.037448	-0.044864	0.055864	-0.042386	0.093094	-0.032761	0.097694	-0.030736
16	0.020031	-0.026016	0.020510	-0.028174	0.062930	-0.027323	0.069028	-0.026363
17	-0.015275	-0.009205	-0.000697	-0.014585	0.038772	-0.020920	0.045309	-0.020974
18	-0.018159	-0.002410	-0.009687	-0.004043	0.021077	-0.014552	0.027143	-0.015393
$\eta=3$								
10	0.21935	-0.069653	0.21519	-0.058326	0.19750	-0.034010	0.19332	-0.030553
11	0.14948	-0.068304	0.15605	-0.058616	0.16199	-0.036452	0.16127	-0.033096
12	0.085936	-0.057454	0.10043	-0.051586	0.12572	-0.035579	0.12809	-0.032814
13	0.036512	-0.040813	0.054523	-0.039660	0.091749	-0.031971	0.096446	-0.030124
14	0.004712	-0.022966	0.021702	-0.025958	0.062393	-0.026524	0.068416	-0.025719
15	-0.010384	-0.011118	0.002215	-0.015514	0.038972	-0.020679	0.045305	-0.020693
16	-0.013000	0.001761	-0.006202	-0.004053	0.021816	-0.014140	0.027616	-0.015006

TABLE 1 (Continued)

ρ	$L=5$		$L=6$		$L=10$		$L=11$	
	f_5	f'_5	f_6	f'_6	f_{10}	f'_{10}	f_{11}	f'_{11}
$\eta=2$								
10	0.082219	-0.053548	0.097149	-0.049049	0.12379	-0.034943	0.12639	-0.032339
11	0.036698	-0.037133	0.053852	-0.037117	0.090545	-0.031188	0.095298	-0.029515
12	0.007982	-0.020616	0.023274	-0.024107	0.061972	-0.025766	0.067890	-0.025102
13	-0.005664	-0.007437	0.005119	-0.012637	0.039236	-0.019672	0.045357	-0.019902
14	-0.008589	0.000703	-0.003060	-0.004320	0.022560	-0.013785	0.028103	-0.014659
$\eta=1$								
10	0.011458	-0.018979	0.025065	-0.022631	0.061649	-0.025055	0.067448	-0.024514
11	-0.001389	-0.007453	0.007878	-0.012181	0.039549	-0.019133	0.045455	-0.019424
12	-0.004991	-0.000522	-0.000309	-0.004743	0.023295	-0.013481	0.028597	-0.014350

The authors wish to thank Dr. F. Hirst and his computer staff for their willing assistance in the operation of CSIRAC and Associate Professor C. B. O. Mohr for his interest in the work.

References

- BLOCH, I., HULL, M. H. JR., BROYLES, A. A., BOURICIUS, W. G., FREEMAN, B. E., and BREIT, G. (1950).—*Phys. Rev.* **80** : 553.
 FROBERG, C. (1955).—*Rev. Mod. Phys.* **27** : 399.
 NATIONAL BUREAU OF STANDARDS (1952).—"Tables of Coulomb Wave Functions." Vol. I. Applied Mathematics Series No. 17.

METHODS FOR NUMERICAL CALCULATIONS WITH THE TYPE I COUNTER

By P. G. GUEST*

[Manuscript received January 20, 1958]

Summary

Detailed schemes have been prepared for performing numerical calculations of frequency, expectation, and variance. Tables are given of the deviations of the asymptotic formulae from the exact formulae when the time of observation is small.

I. INTRODUCTION

A number of papers (Jost 1947 ; Feller 1948 ; Elmore 1950 ; Ramakrishnan and Mathews 1953 ; Hull and Wolfe 1954 ; Campbell 1956 ; Takacs 1956) dealing with the Type I counter have appeared in recent years. The present paper differs from these in being concerned primarily with the calculation of numerical magnitudes rather than with the derivation of general formulae.

A Type I counter is one in which each count is followed by a fixed dead time τ during which the counter is insensitive, the counter returning to the sensitive condition at the end of the period τ . Any event occurring in the dead time has no effect. It will be assumed that the events which are being counted are distributed in time according to a Poisson distribution with parameter μ . The quantities which are of interest are the frequency function $F(t)$ and the expected number $N(t)$ of counts in time t . $F(t)dt$ is the probability that a count occurs in the interval $(t, t+dt)$ and

$$N(t) = \int_0^t F(x)dx.$$

The counting will be assumed to start at $t=0$, the counter being sensitive at that instant. $F(t)$ has the form shown in Figure 1, which corresponds to the special case $\mu\tau=1$. At $t=0$, $F(t)$ equals μ , and $F(t)$ oscillates with decreasing amplitude about the value $\mu/(1+\mu\tau)$ as t increases. When t is large $F(t)$ may be assumed to have this value and $N(t)$ may be taken as $\mu t/(1+\mu\tau)$. If the events occur in short bursts, as will happen when particles produced by a pulsed source are being counted, then the behaviour of $F(t)$ and $N(t)$ must be investigated for small values of t .

Instead of supposing the counter to be sensitive at $t=0$, it is sometimes convenient to consider the case where a count occurs at $t=0$. Clearly, in this case the counter will be sensitive at $t=\tau$, and the frequency, expectation, and variance of subsequent counts will be given by the same formulae as for the case where the counter is sensitive at $t=0$ if t in these formulae is replaced by $t-\tau$.

* Physics Department, University of Sydney.

II. EXACT VALUES OF FREQUENCY AND EXPECTATION

The function $F(t)$ is given by the expression

$$F(t) = \mu \sum_{k=0}^p e^{-\mu(t-k\tau)} \frac{\mu^k (t-k\tau)^k}{k!}, \dots\dots\dots (1)$$

where

$$t = p\tau + \alpha, \quad 0 \leq \alpha < \tau. \dots\dots\dots (2)$$

The k th term of the sum gives, when multiplied by dt , the probability that the $(k+1)$ th count occurs in $(t, t+dt)$. The term differs from the corresponding term for a Poisson distribution in that the dead times following each of the first k counts are subtracted from t .

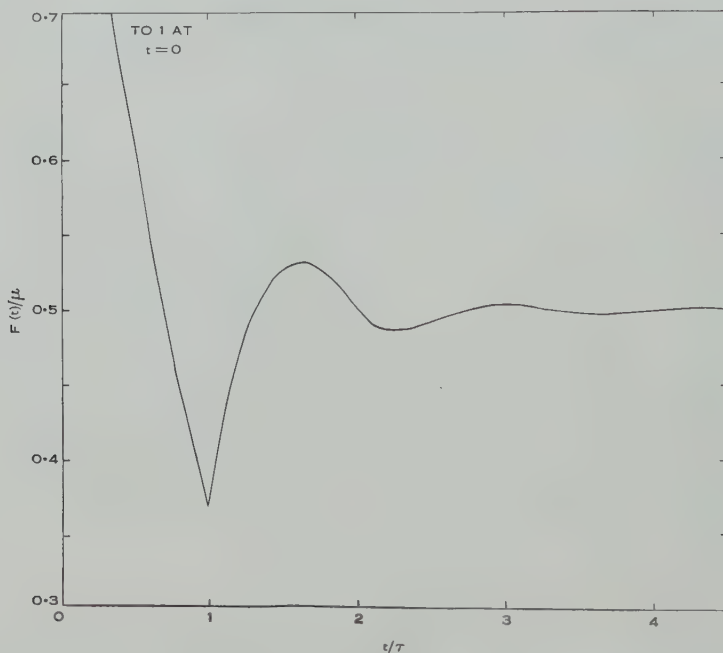


Fig. 1.—Graph of $F(t)/\mu$ against t/τ for a Type I counter with $\mu\tau=1$.

$N(t)$ may be found by integrating $F(t)$, giving

$$N(t) = \sum_{k=0}^p \Gamma(k+1; \mu t - k\mu\tau) / \Gamma(k+1), \dots\dots\dots (3)$$

the terms $\Gamma(k+1; x)$ being the incomplete gamma functions. Alternatively, the equation

$$F(t) = \mu \left\{ 1 - \int_{t-\tau}^t F(x) dx \right\} \dots\dots\dots (4)$$

may be used, the term in brackets being the probability that the counter is sensitive at the instant t (i.e. that no count has occurred in $(t-\tau, t)$). (4) can be written

$$F(t) = \mu \{ 1 - N(t) + N(t-\tau) \},$$

giving the recurrence relation

$$N(t) = 1 + N(t - \tau) - \mu^{-1} F(t),$$

and so

$$N(t) = p + 1 - \mu^{-1} \sum_{j=0}^p F(t - j\tau),$$

or, using (1),

$$N(t) = p + 1 - \sum_{k=0}^p e^{-\mu(t-k\tau)} \exp_k \mu(t-k\tau), \quad \dots\dots\dots (5)$$

where

$$\exp_k t = \sum_{j=0}^k t^j / j! \quad \dots\dots\dots (6)$$

Formulae more suitable for numerical calculations are obtained in terms of the variable

$$\beta = \mu(t - p\tau) = \mu\alpha. \quad \dots\dots\dots (7)$$

It is found that

$$F_p(\beta) = \mu e^{-\beta} \sum_{k=0}^p a_{p-k} \beta^k / k! \quad \dots\dots\dots (8)$$

and

$$N_p(\beta) = p + 1 - e^{-\beta} \sum_{k=0}^p b_{p-k} \beta^k / k! \quad \dots\dots\dots (9)$$

the coefficients a_r and b_r being functions of $\mu\tau$ given by the expressions

$$a_r = \sum_{j=0}^r (j\mu\tau)^{r-j} e^{-j\mu\tau} / (r-j)! \quad \dots\dots\dots (10)$$

and

$$b_r = \sum_{s=0}^r a_s = b_{r-1} + a_r. \quad \dots\dots\dots (11)$$

TABLE 1
SCHEME FOR THE CALCULATION OF a_r AND b_r

A r	1	$e^{-\mu\tau}$	$e^{-2\mu\tau}$	$e^{-3\mu\tau}$	$e^{-4\mu\tau}$	a_r	b_r
0	1					a_0	$b_0 = a_0$
1		1				a_1	$b_1 = b_0 + a_1$
2		$\mu\tau$	1			a_2	$b_2 = b_1 + a_2$
3		$\frac{1}{2}(\mu\tau)^2$	$2\mu\tau$	1		a_3	$b_3 = b_2 + a_3$
4		$\frac{1}{6}(\mu\tau)^3$	$\frac{1}{2}(2\mu\tau)^2$	$3\mu\tau$	1	a_4	$b_4 = b_3 + a_4$

A scheme for the calculation of a_r and b_r is shown in Table 1. The entries in each column are the successive terms of $e^{j\mu\tau}$. a_r is the sum of the products of corresponding elements in row r and row A . The sums in (8) and (9) can be evaluated by successive multiplications and additions. For example, when p is 3,

$$\sum a_{p-k} \beta^k / k! = \{(0.16\beta + 0.5a_1)\beta + a_2\}\beta + a_3.$$

Example 1

What is the value of $N(t)$ when t equals 450 μsec , if τ is 134 μsec and μ is 4070 events per second?

Here $\mu\tau$ is 0.545380 and μt is 1.831500, and so p is 3 and β is 0.195360. The values a_r and b_r are calculated in Table 2. From formula (9)

$$N_p(\beta) = 4 - 3 \cdot 346455e^{-\beta} = 1.247411.$$

TABLE 2
CALCULATION OF a_r AND b_r FOR EXAMPLE 1

A r	1	0.579622	0.335962	0.194731	a_r	b_r
0	1				1	1
1		1			0.579622	1.579622
2		0.545380	1		0.652076	2.231698
3		0.148720	1.090760	1	0.647386	2.879084
Sum, S	1	1.694100	2.090760	1	Check, $S \times A$ 2.879085	

For values of p less than 2 the formulae for $F(t)$ and $N(t)$ are :

$$\begin{aligned} p=0: & \quad \mu^{-1}F(t) = e^{-\beta}, & N(t) &= 1 - e^{-\beta}, \\ p=1: & \quad \mu^{-1}F(t) = e^{-\beta}(\beta + e^{-\mu\tau}), & N(t) &= 2 - e^{-\beta}(1 + \beta + e^{-\mu\tau}). \end{aligned}$$

For higher values of p it is best to use the scheme of Table 1.

III. ASYMPTOTIC FORMULAE

Equation (4) can be integrated to give

$$N(t) = \mu t - \mu \int_{t-\tau}^t N(x) dx. \quad \dots\dots\dots (12)$$

If the asymptotic formula, valid for large t , is written

$$n(t) = \gamma(\mu t + \psi), \quad \dots\dots\dots (13)$$

γ and ψ can be found by substituting this expression in (12) and equating terms in t on each side and constant terms on each side. This gives

$$\gamma = 1/(1 + \mu\tau), \quad \psi = \frac{1}{2}\gamma(\mu\tau)^2. \quad \dots\dots\dots (14)$$

The asymptotic value $f(t)$ of $F(t)$ is of course equal to γ .

For Example 1,

$$\gamma = 0.647090, \quad \gamma\psi = 0.062273, \quad \mu t = 1.831500,$$

and

$$n(t) = 1.185145 + 0.062273 = 1.247418.$$

This is very close to the exact value. Since the asymptotic values are easily calculated, it is of interest to determine by how much the asymptotic values differ from the exact values when μt is small. This question has been investigated by determining the maximum values of the fractional divergences

$$|F(t) - f(t)|/F(t) \text{ and } |N(t) - n(t)|/N(t)$$

for values of t greater than $p\tau$, $p=1(1)4$. The maximum divergences (expressed as percentages) of $f(t)$ are shown in Table 3 and of $n(t)$ in Table 4 for values of $\mu\tau$ up to 2. From these tables it is possible to determine in any particular problem whether the asymptotic formula is sufficiently accurate.

TABLE 3
MAXIMUM PERCENTAGE DEVIATION OF ASYMPTOTIC VALUE $f(t)$ FROM
EXACT VALUE $F(t)$ WHEN $t \geq p\tau$

p $\mu\tau$	1	2	3	4
0.0125	0.01			
0.025	0.03			
0.05	0.12			
0.075	0.27	0.01		
0.1	0.5	0.01		
0.15	1.0	0.04		
0.2	1.8	0.09		
0.3	3.8	0.24		
0.4	7	0.4	0.02	
0.5	10	0.7	0.05	0.01
0.6	14	0.9	0.11	0.02
0.7	18	1.0	0.21	0.04
0.8	24	1.2	0.37	0.05
0.9	29	1.8	0.6	0.07
1.0	36	2.5	0.9	0.10
1.2	51	4	1.6	0.27
1.4	69	7	2.4	0.6
1.6	91	11	3.1	1.0
1.8	116	15	3.6	1.7
2.0	146	21	5	2.6

For Example 1, $\mu\tau$ is 0.55 and p is 3, and so $n(t)$ should differ from $N(t)$ by an amount not exceeding 0.01 per cent.

It is shown in Section IV that the standard deviation is approximately equal to $\sqrt{(\gamma^3 \mu t)}$, and so the ratio of the constant term $\gamma\psi$ in $n(t)$ to the standard deviation is $\psi/\sqrt{(\gamma \mu t)}$, which will be considerably less than unity unless t is very

small. This term can then often be omitted and $n(t)$ taken as simply $\gamma\mu t$. However, when the experiment consists of a large number of repeated determinations over the time t it is advisable to retain the constant term to avoid bias.

TABLE 4
MAXIMUM PERCENTAGE DEVIATION OF ASYMPTOTIC VALUE $n(t)$ FROM
EXACT VALUE $N(t)$ WHEN $t \geq p\tau$

p $\mu\tau$	1	2	3	4
0.025	0.01			
0.05	0.04			
0.075	0.08			
0.1	0.13			
0.15	0.25			
0.2	0.39	0.01		
0.3	0.7	0.01		
0.4	1.0	0.02		
0.5	1.2	0.04	0.01	
0.6	1.3	0.07	0.01	
0.7	1.4	0.12	0.02	
0.8	1.4	0.18	0.02	
1.0	1.4	0.35	0.03	0.01
1.2	2.1	0.6	0.06	0.02
1.4	2.8	0.8	0.10	0.04
1.6	3.7	1.0	0.17	0.05
1.8	4.6	1.2	0.26	0.06
2.0	5.6	1.3	0.37	0.08

IV. VARIANCE

The variance of the observed number of counts is given by the formula

$$V(t) = \sum_{k=0}^p \{2(p-k)+1\} e^{-\mu(t-k\tau)} \exp_k \mu(t-k\tau) - \left\{ \sum_{k=0}^p e^{-\mu(t-k\tau)} \exp_k \mu(t-k\tau) \right\}^2. \quad (15)$$

For purposes of numerical calculation it is convenient to introduce the variable β defined by equation (7), and to write

$$V_p(\beta) = e^{-\beta} \sum_{j=0}^p g_{p-j} \beta^j / j! - \{ e^{-\beta} \sum_{j=0}^p b_{p-j} \beta^j / j! \}^2, \quad (16)$$

where

$$g_r = \sum_{s=0}^r f_s \quad \text{and} \quad f_r = \sum_{j=0}^r (2j+1) e^{-j\mu\tau} (j\mu\tau)^{r-j} / (r-j)! \quad (17)$$

It will be observed that the terms of f_r differ from those of a_r only by the factor $(2j+1)$, and so the scheme of Table 1 can be used to calculate f_r and g_r if the row A is replaced by a row F whose elements are $(2j+1)e^{-j\mu\tau}$.

Table 5 shows how the calculations given in Table 2 for Example 1 may be extended to include the variance calculations. Using (16), with $\beta=0.195360$,

$$V_3(\beta)=9.923019e^{-\beta}-(3.346455e^{-\beta})^2=0.585318.$$

For values of p less than 2 the formulae for $V(t)$ are :

$$\begin{aligned} p=0 &: e^{-\beta}-e^{-2\beta}, \\ p=1 &: e^{-\beta}(1+\beta+3e^{-\mu\tau})-\{e^{-\beta}(1+\beta+e^{-\mu\tau})\}^2. \end{aligned}$$

For higher values of p it is best to use the scheme of Table 5.

TABLE 5
EXTENSION OF TABLE 2 TO PERMIT THE CALCULATION OF VARIANCE

F A r	1	1.738866	1.679810	1.363117				
	1	0.579622	0.335962	0.194731	a_r	b_r	f_r	g_r
0	1				1	1	1	1
1		1			0.579622	1.579622	1.738866	2.738866
2		0.545380	1		0.652076	2.231698	2.628153	5.367019
3		0.148720	1.090760	1	0.647386	2.879084	3.453991	8.821010
Sum, S	1	1.694100	2.090760	1	Check, $A \times S$	2.879085	Check, $F \times S$	8.821009

The asymptotic formula $v(t)$ for the variance is

$$v(t)=\gamma^3(\mu t+\chi), \quad \dots\dots\dots (18)$$

where

$$\gamma=1/(1+\mu\tau), \quad \chi=\frac{1}{12} \gamma(\mu\tau)^2\{18+4\mu\tau+(\mu\tau)^2\}. \quad \dots\dots (19)$$

For Example 1,

$$\gamma=0.647090, \quad \gamma^3=0.270953, \quad \chi=0.328465, \quad \mu t=1.831500,$$

and so

$$v(t)=0.585249.$$

This is very close to the value given by the exact formula. Calculations have been made of the divergence of the asymptotic value $v(t)$ from the exact value $V(t)$ when μt is small. Since the variance is seldom required very accurately, a detailed table similar to Table 4 will not be given. It is found that the percentage difference between the two values is less than 10 per cent. for all $t \geq \tau$ when $\mu\tau < 0.89$, and for all $t \geq 2\tau$ when $\mu\tau < 1.78$.

For small values of μt the standard deviation (the square root of the variance) will be of the same order as the expectation. However, if the average number of counts is taken over a large number k of intervals of duration t , the standard deviation will be divided by the factor \sqrt{k} .

V. ESTIMATION OF μ

In many investigations an experimental determination n of the counts in time t is made and an estimate of the parameter μ is required. The value μ_0 derived from the asymptotic formulae (13) and (14) is

$$\mu_0 = \frac{\sqrt{(t^2 + 2n\tau^2)} - (t - 2n\tau)}{2t\tau - (2n - 1)\tau^2}. \quad \dots\dots\dots (20)$$

Example 2

For a counter with dead time 134 μ sec the average number of counts in 350 μ sec is 1.276. What is the value of μ ?

From (20),

$$\mu_0 = \frac{410.27 - 8.03}{(93800 - 27868) \times 10^{-6}} = 6100.8.$$

Since p is 2 and $\mu_0\tau$ is 0.82, the error through use of the asymptotic formula is, from Table 4, less than 0.2 per cent.

TABLE 6
CALCULATION OF VALUE OF $N(t)$ CORRESPONDING TO μ_0 (EXAMPLE 2)
 $\mu_0\tau = 0.817507$, $\mu_0t = 2.135280$, $p = 2$, $\beta = 0.500266$

C	0	0.441531	0.389900			
A	1	0.441531	0.194950			
r				a_r	b_r	c_r
0	1			1	1	0
1		1		0.441531	1.441531	0.441531
2		0.817507	1	0.555905	1.997436	0.750855

$$N(t) = 3 - 2.843718e^{-\beta} = 1.275656$$

As a check, the value of $N(t)$ corresponding to μ_0 is calculated in Table 6. This differs from n by 0.000344, which is negligible in the present case. However, if it is desired to obtain the value μ for which $N(t)$ is exactly equal to n , the formula

$$\mu \frac{dN(t)}{d\mu} = e^{-\beta} (\beta \sum a_{p-j} \beta^j / j! + \mu \tau \sum c_{p-j} \beta^j / j!) \quad \dots\dots\dots (21)$$

may be used. The quantities c_r are calculated in the same way as the quantities a_r , with $e^{-j\mu\tau}$ replaced by $j e^{-j\mu\tau}$. In Table 6 the c_r are calculated for Example 2 by intermultiplying row C and row r . From (21), the change $d\mu$ required to increase $N(t)$ by 0.000344 is given by

$$\begin{aligned} 0.000344 &= (d\mu/\mu_0) e^{-\beta} (\beta \sum a_{p-j} \beta^j / j! + \mu \tau \sum c_{p-j} \beta^j / j!) \\ &= (d\mu/\mu_0) 0.606370 (0.451200 + 0.794403), \end{aligned}$$

and so

$$\mu/\mu_0 = (\mu_0 + d\mu)/\mu_0 = 1.000455,$$

and

$$\mu = 6103.58.$$

A recalculation of Table 6 with this new value of μ does in fact give $N(t)=1.276000$.

By differentiation of (20) it is found that

$$\frac{d\mu_0}{dn} = \frac{1}{t} \left\{ \frac{1 + \mu_0\tau + \tau/2 \sqrt{(t^2 + 2n\tau^2)}}{1 - (2n-1)\tau/2t} \right\}, \dots\dots\dots (22)$$

and an estimate of the variance of μ_0 may be obtained from

$$\text{var } \mu_0 = (d\mu_0/dn)^2 \text{var } n. \dots\dots\dots (23)$$

For Example 2, $d\mu_0/dn$ is 8052. To obtain an estimate of $\text{var } n$, the approximate value μ_0 may be substituted in equations (18) and (19), leading to the value 0.467622. The standard deviation of μ_0 is then 5506. If the estimate n is the average of k determinations, this quantity should be divided by \sqrt{k} .

VI. SCALING CIRCUITS

The probability $Q(k+1, t)$ that $k+1$ or more counts occur in time t is given by the k th term of the expansion (5) of $N(t)$,

$$Q(k+1, t) = 1 - e^{-\mu(t-k\tau)} \exp_k \mu(t-k\tau). \dots\dots\dots (24)$$

This differs from the corresponding formula for the Poisson distribution in that t is replaced by $t-k\tau$. Molina (1942) in his Table II tabulates a sum $P(c, a)$ which may be used to obtain $Q(k, t)$, the relation being

$$Q(k, t) = P(c=k; a=\mu t - (k-1)\mu\tau). \dots\dots\dots (25)$$

The square root approximation to the significance limits of the Poisson distribution (Blom 1954) leads to the formula

$$\sqrt{k} - \sqrt{\{\mu t - (k-1)\mu\tau\}} = \frac{1}{2}X, \dots\dots\dots (26)$$

where X is the value of the standardized normal variate corresponding to the significance level Q .

Since mechanical recorders have dead times τ^* of the order of 0.1 sec, it is often necessary to interpose a scaling circuit between the counter and the recorder to reduce the number of counts reaching the recorder by a factor m . If more than m counts occur in the time $\tau^* - \tau$ during which the counter has become sensitive again following a recorded count while the recorder is still insensitive, there will be a counting loss in the recorder. It is customary to choose the scaling factor m so that this loss can be neglected. This means choosing m so that $Q(m, \tau^* - \tau)$ is negligible, using either Molina's table with

$$Q(m, \tau^* - \tau) = P(m, \mu\tau^* - m\mu\tau), \dots\dots\dots (27)$$

or the square root approximation

$$\sqrt{m} - \sqrt{(\mu\tau^* - m\mu\tau)} = \frac{1}{2}X. \dots\dots\dots (28)$$

When m is small the estimate from (28) will be slightly higher than the exact estimate from (27).

Example 3

If τ is 134 μ sec and μ does not exceed 4000 events per second, what scaling factor is required in front of a fast mechanical recorder of dead time 0.02 sec if the scaling loss is to be less than 0.1 per cent.?

The values $\mu\tau^*$ and $\mu\tau$ are 80 and 0.536. Molina's table gives

$$P(66, 44.6) = 0.0017, \quad P(67, 44.1) = 0.0008,$$

and so m must exceed 66 if the scaling losses are to be less than 0.1 per cent. For the square root approximation, X is 3.09 for the 0.1 per cent. level, while

$$\sqrt{66} - \sqrt{44.6} = 1.45, \quad \sqrt{67} - \sqrt{44.1} = 1.54, \quad \sqrt{68} - \sqrt{43.6} = 1.64,$$

which would imply that m must be greater than 67.

VII. TWO COUNTERS IN SERIES

If the functions for the second counter are distinguished by an asterisk,

$$F^*(t) = F(t) - \int_{\tau}^{\tau^*} F^*(t-x)F(x-\tau)dx, \quad \dots\dots\dots (29)$$

and

$$N^*(t) = N(t) - \int_{\tau}^{\tau^*} N^*(t-x)F(x-\tau)dx. \quad \dots\dots\dots (30)$$

The exact expressions for $F^*(t)$ and $N^*(t)$ will be quite complicated, but the asymptotic forms can be found fairly easily. It will be observed first that, when $\tau^* \leq \tau$, $N^*(t)$ is equal to $N(t)$ and the second counter has no effect on the counting rate. This condition should clearly be aimed at in the design of the equipment.

To find the asymptotic formula $n^*(t)$ for $N^*(t)$ when $\tau^* > \tau$,

$$n^*(t) = \gamma^*(\mu t + \psi^*) \quad \text{and} \quad n(t) = \gamma(\mu t + \psi)$$

are substituted in (30), and terms in t equated and constant terms equated. This gives

$$\gamma^* = \gamma \{1 + N(\tau^* - \tau)\}^{-1}, \quad \psi^* = \psi + (\gamma^*/\gamma) \int_{\tau}^{\tau^*} x \mu F(x - \tau) dx. \quad \dots (31)$$

The integral can be put in the form

$$\int_{\tau}^{\tau^*} x \mu F(x - \tau) dx = \frac{1}{2} q(q+1)(1 + \mu\tau) - \sum_{p=0}^{q-1} (\mu\tau^* + q - p) \mu^{-1} F_p(\beta^*), \quad \dots (32)$$

where

$$\mu\tau^* = q\mu\tau + \beta^*.$$

Example 4

A Geiger counter with a dead time of 100 μ sec is operating a recording circuit whose dead time is 240 μ sec. If the rate of arrival of the particles being counted is 1000 per second, find an expression for the average number of counts in time t .

In this case

$$\begin{aligned}\mu\tau &= 0.1, & \mu\tau^* &= 0.24, & q &= 2, & \beta^* &= 0.04, \\ N(\tau^* - \tau) &= N_1(\beta^*) & &= 0.131422, \\ \mu^{-1}F_0(\beta^*) &= 0.960789, & \mu^{-1}F_1(\beta^*) &= 0.907789, \\ \gamma &= 0.909091, & \psi &= 0.004545.\end{aligned}$$

Hence, from (31) and (32),

$$\gamma^* = 0.803494, \quad \psi^* = 0.004545 + 0.022174\gamma^*/\gamma = 0.024143,$$

and the average number $n^*(t)$ of particles counted in time t will be given by

$$n^*(t) = 0.803494(1000t + 0.024143).$$

No detailed investigation of the deviations of the asymptotic values from the exact values has been made for the case of two counters in series, but it seems probable that the percentage deviations are similar to those for a single counter of dead time τ^* . The variance of the number of counts is quite close to $\gamma^{*3}\mu t$ when μt is large; in fact,

$$v^*(t) = \gamma^{*3}(\kappa\mu t + \chi^*), \quad \dots\dots\dots (33)$$

where

$$\kappa = \{2\gamma^*(1 + \psi^*) - 1\}/\gamma^{*2}$$

is very close to unity. The expression for the constant term χ^* is very complicated.

VIII. REFERENCES

- BLOM, G. (1954).—Transformations of the binomial, negative binomial, Poisson and χ^2 distributions. *Biometrika* **41**: 302.
- CAMPBELL, L. L. (1956).—Standard deviation of dead time correction in counters. *Canad. J. Phys.* **34**: 929.
- ELMORE, W. C. (1950).—Statistics of counting. *Nucleonics* **6**: 26.
- FELLER, W. (1948).—On probability problems in the theory of counters. Courant Anniversary Volume, p. 105. (Interscience Publishers: New York.)
- HULL, T. E., and WOLFE, W. A. (1954).—On inverting Laplace transforms. *Canad. J. Phys.* **32**: 72.
- JOST, R. (1947).—Bemerkungen zur mathematischen Theorie der Zähler. *Helv. Phys. Acta* **20**: 173.
- MOLINA, E. C. (1942).—"Poisson's Exponential Binomial Limit." (D. van Nostrand Co.: New York.)
- RAMAKRISHNAN, A., and MATHEWS, P. M. (1953).—A stochastic problem relating to counters. *Phil. Mag.* **44**: 1122.
- TAKACS, L. (1956).—On a probability problem arising in the theory of counters. *Proc. Camb. Phil. Soc.* **52**: 488.

THE GENERAL RELATIONSHIPS BETWEEN THE ELASTIC CONSTANTS OF ISOTROPIC MATERIALS IN n DIMENSIONS

By N. W. TSCHOEGL*

[Manuscript received December 30, 1957]

Summary

The relationships between the elastic constants of homogeneous isotropic materials in n dimension are derived and are shown to depend on n . The maximal value of the generalized Poisson's ratio is $1/(n-1)$. The n -dimensional formulae reproduce the well-known three-dimensional relations for $n=3$, while $n=2$ produces the relations appropriate for monomolecular films. The correct degeneration is shown for $n=1$.

I. INTRODUCTION

In the physical chemistry of monomolecular films it is customary to regard the film (*or monolayer*) as a two-dimensional system since there is no bonding between molecules of the same kind in the third dimension. One accordingly defines a *surface bulk (compressional) modulus* and a *surface shear modulus* having the dimensions of surface traction (force per unit length) instead of traction (force per unit area) (Langmuir and Schaefer 1937).† This procedure prompts the development of a general n -dimensional elastic theory comprising both two- and three-dimensional theory as special cases.

In this paper it is proposed to derive the general relationships connecting the four elastic material constants of homogeneous isotropic bodies in n dimensions. It will be shown that these relations are generally dependent on the number of dimensions considered.

II. THE FUNDAMENTAL ELASTIC MODULI IN n DIMENSIONS

An infinitesimal elastic deformation of a homogeneous isotropic body can be resolved into a change in size (dilatation or contraction) and an independent change in shape (distortion). Homogeneous isotropic bodies therefore possess two fundamental elastic moduli, the *bulk modulus* relating to changes in size and the *shear modulus* relating to changes in shape.

It follows from the independence of the two fundamental moduli that the general n -dimensional stress tensor s_{ij} is not proportional to the strain tensor e_{ij} . These tensors may, however, be resolved into separately proportional component tensors such that

$$s''_{ij} = s_{ij} - s'_{ij}, \quad \dots\dots\dots (1a)$$

$$e''_{ij} = e_{ij} - e'_{ij}, \quad \dots\dots\dots (1b)$$

* Bread Research Institute of Australia, North Sydney, N.S.W.

† LANGMUIR, I., and SCHAEFER, N. J. (1937).—*J. Amer. Chem. Soc.* **59**: 2400.

where s'_{ij} and e'_{ij} may be called *mean normal tensors* while s''_{ij} and e''_{ij} are termed *deviators*.

We can now state

$$s'_{ij} = nK_n e'_{ij}, \quad \dots \quad (2a)$$

$$s''_{ij} = 2G_n e''_{ij}, \quad \dots \quad (2b)$$

where K_n and G_n are the n -dimensional bulk modulus and shear modulus respectively.

The dimensional factor in (2a) appears because K_n is traditionally defined as the proportionality constant between the mean normal stress s_{ii}/n and the (hyper)volumetric strain e_{ii} ,

$$s_{ii}/n = K_n e_{ii}. \quad \dots \quad (2a')$$

That (2a) and (2a') are equivalent may be seen from the following development.

Since the deviatoric tensors relate to changes in shape only, the sums of their principal components vanish. Using the summation convention

$$s''_{ii} \equiv 0, \quad \dots \quad (3a)$$

$$e''_{ii} \equiv 0, \quad \dots \quad (3b)$$

and, from (1a) and (1b)

$$s'_{ii} = s_{ii}, \quad \dots \quad (4a)$$

$$e'_{ii} = e_{ii}. \quad \dots \quad (4b)$$

The mean normal tensors relate to changes in size only. Clearly their non-principal components must vanish while their principal components must be equal. It then follows from (4a) and (4b) that

$$s'_{ij} = s_{ii}/n \cdot d_{ij}, \quad \dots \quad (5a)$$

$$e'_{ij} = e_{ii}/n \cdot d_{ij}, \quad \dots \quad (5b)$$

where d_{ij} is the unit tensor. Substituting (5a) and (5b) into (2a) then leads to (2a').

The numerical factor 2 in (2b) arises because of the traditional definition of shear which is twice the corresponding shear component of the strain tensor.

III. THE RELATIONSHIPS BETWEEN THE ELASTIC CONSTANTS

To derive the general relationships between K_n , G_n , the n -dimensional *extensional* (Young's) modulus Y_n , and *Poisson's ratio* μ_n we consider a pure tensile stress in the n th dimension. For such a stress all components of the stress tensor vanish except s_{nn} . All non-principal components of the strain tensor also vanish since a pure tensile stress produces no distortion. The extension e_{nn} in the n th dimension is accompanied by $(n-1)$ contractions $-\mu_n e_{nn}$ in the remaining dimensions. Consequently

$$s_{ii} = s_{nn}, \quad \dots \quad (6a)$$

$$e_{ii} = e_{nn} - (n-1)\mu_n e_{nn}. \quad \dots \quad (6b)$$

Substituting into (5a) and (5b) and introducing Y_n from Hooke's law, the mean normal tensors become

$$s'_{ij} = s_{nn}/n \cdot d_{ij}, \quad \dots\dots\dots (7a)$$

$$e'_{ij} = [1 - (n-1)\mu_n]s_{nn}/nY_n \cdot d_{ij}. \quad \dots\dots\dots (7b)$$

The non-principal components of the deviators also vanish since there is no distortion. The principal components can be expressed by the following set of equations

$$s''_{11} = s''_{22} = \dots = s''_{(n-1)(n-1)} = 0 - s_{ii}/n, \quad \dots\dots\dots (8a)$$

$$s''_{nn} = s_{nn} - s_{ii}/n, \quad \dots\dots\dots (8a')$$

$$e''_{11} = e''_{22} = \dots = e''_{(n-1)(n-1)} = -[\mu_n e_{nn} - e_{ii}/n], \quad \dots\dots\dots (8b)$$

$$e''_{nn} = e_{nn} - e_{ii}/n, \quad \dots\dots\dots (8b')$$

obtained by subtracting the principal components of the mean normal tensors from those of the total tensors according to (1a) and (1b).

Substituting for s_{ii} from (6a) and for e_{ii} from (6b) and using Hooke's law

$$s''_{11} = s''_{22} = \dots = s''_{(n-1)(n-1)} = -s_{nn}/n, \quad \dots\dots\dots (9a)$$

$$s''_{nn} = (n-1)s_{nn}/n, \quad \dots\dots\dots (9a')$$

$$e''_{11} = e''_{22} = \dots = e''_{(n-1)(n-1)} = -(1 + \mu_n)s_{nn}/nY_n, \quad \dots\dots\dots (9b)$$

$$e''_{nn} = (n-1)(1 + \mu_n)s_{nn}/nY_n. \quad \dots\dots\dots (9b')$$

Defining a tensor g_{ij} such that

$$g_{11} = g_{22} = \dots = g_{(n-1)(n-1)} = -1, \quad \dots\dots\dots (10)$$

$$g_{nn} = n-1, \quad \dots\dots\dots (10')$$

while all non-principal components vanish, the deviatoric tensors may be written

$$s''_{ij} = s_{nn}/n \cdot g_{ij}, \quad \dots\dots\dots (11a)$$

$$e''_{ij} = (1 + \mu_n)s_{nn}/nY_n \cdot g_{ij}. \quad \dots\dots\dots (11b)$$

Now from (2a), (7a), and (7b)

$$K_n = \frac{Y_n}{n[1 - (n-1)\mu_n]}, \quad \dots\dots\dots (12)$$

and from (2b), (11a), and (11b)

$$G_n = \frac{Y_n}{2(1 + \mu_n)}. \quad \dots\dots\dots (13)$$

From (12) and (13)

$$Y_n = \frac{2n^2 K_n G_n}{n(n-1)K_n + 2G_n}, \quad \dots\dots\dots (14)$$

$$\mu_n = \frac{nK_n - 2G_n}{n(n-1)K_n + 2G_n}. \quad \dots\dots\dots (15)$$

IV. THE MAXIMAL VALUE OF POISSON'S RATIO

Writing (15) in the form

$$nK_n[1-(n-1)\mu_n]=2G_n(1+\mu_n), \quad \dots\dots\dots (15')$$

it is apparent that μ_n cannot be less than -1 and cannot exceed $1/(n-1)$. The maximal value of Poisson's ratio, $\mu_{n(\max)}$, is thus seen to depend on n .

The identical result can also be obtained directly by the following consideration: let an n -dimensional hypercube of unit length be extended by a small amount λ in one dimension. The change ΔH in the hypervolume then is

$$\Delta H=(1+\lambda)(1-\mu_n\lambda)^{(n-1)}-1. \quad \dots\dots\dots (16)$$

Expanding $(1-\mu_n\lambda)^{(n-1)}$ by the binomial theorem and neglecting all higher powers of λ

$$\Delta H=(1+\lambda)[1-(n-1)\mu_n\lambda]-1. \quad \dots\dots\dots (17)$$

If the hypercube is incompressible ($K_n=\infty$), $\Delta H=0$, and again neglecting λ^2 ,

$$\mu_{n(\max)}=1/(n-1). \quad \dots\dots\dots (18)$$

Either from (13) by substituting $\mu_n=1/(n-1)$, or from (14) by substituting $K_n=\infty$, we obtain the general relationships between the extensional and shear moduli of incompressible bodies (for $n>1$) as

$$Y_n=\frac{2n}{(n-1)}G_n. \quad \dots\dots\dots (19)$$

V. SPECIAL CASES

For $n=3$, (12)-(15) and (18) and (19) reproduce the well-known three-dimensional relationships. The two-dimensional relations can be obtained simply by substituting $n=2$ and are seen to be different from the relations in three dimensions. For $n=1$, (12) and (14) correctly show that the bulk modulus of a "one-dimensional body" degenerates into the extensional modulus

$$K_1\equiv Y_1. \quad \dots\dots\dots (20)$$

This modulus is the only elastic constant a "one-dimensional body" can have. Neither G_n nor μ_n have any meaning for $n=1$. In fact, expressing G_1 and μ_1 in terms of K_1 and Y_1 from (14) and (12) and using identity (20) we obtain, for $n=1$,

$$G_1=\frac{n(n-1)K_1}{2(n-1)(n+1)}=0/0, \quad \dots\dots\dots (21)$$

$$\mu_1=\frac{n-1}{n-1}=0/0. \quad \dots\dots\dots (22)$$

THE IONIZATION OF LIQUID ARGON BY α -PARTICLES

By F. D. STACEY*

[*Manuscript received February 19, 1958*]

Summary

Calculations suggest that in liquid argon the column of ionization produced by an α -particle consists of a core of positive charge $0.6-1.5 \times 10^{-5}$ cm in diameter concentric with a larger cloud of electrons extending to a diameter of 40×10^{-5} cm. This leads to a new description of charge separation in an applied electric field, in which the positive and negative charges are drawn apart slightly (polarized) and the more energetic electrons escape from the edge of the column. The remaining electron cloud reaches a state of dynamic equilibrium with the excess positive charge in a time short compared with that required for ionic recombination, which does not have a significant effect on the ionization current. The distribution of electron energies in the column of ionization is calculated from the results of recent ionization current measurements, and indicates that a 5.3 MeV α -particle expends 0.5 MeV as kinetic energy of the liberated electrons.

I. INTRODUCTION

Measurements of ionization currents in α -irradiated liquid rare gases (Williams and Stacey 1957) showed that the currents were too large to be explained by either of the existing theories of ionic recombination (Jaffé 1913; Kramers 1952). Both of these theories assume that the electrons released by ionization form heavy negative ions by attachment to neutral atoms, immediately they are formed. This is a questionable assumption for any material, but in liquid argon it is quite certain that the electrons remain free altogether (Williams 1957). Liquid argon is therefore probably the simplest material to re-examine theoretically.

The theories of Jaffé (1913) and Kramers (1952) describe columns of positive and negative ions which have similar mobilities and which, in the absence of an applied field, remain completely superimposed. The application of a field to such a system was considered to draw the two opposite charge clouds apart at a rate dependent upon the field, while ionic recombination removed ions of both signs from those parts of the charge clouds which overlapped. The increase of ionization current with field was thus explained as the reduced opportunity for recombination when the two charge clouds were drawn apart more rapidly, and the general features of the ionization current versus field curve were explained for a number of materials.

Gerritsen (1945), who experimented with a number of liquids, including argon, and who was responsible for the posthumous publication of Kramers' theory, maintained that this theory fitted his data for liquid argon reasonably well. It appears, however, that neglect of the mutual attraction of the opposite

* Department of Geophysics, Australian National University, Canberra.

charge clouds makes the theory quite invalid in the case of liquid argon, and possibly for most other materials.

Considering only the charge collected from a single α -particle track in a small field and neglecting the charge which is lost in ionic recombination, we see from the experimental data (Williams and Stacey 1957) that about 7 ion pairs are collected per volt/cm of applied field; i.e. in a field of F V/cm, $=V$ e.s.u., a total charge of $(7eF)$ is collected per α -particle. The charges are drawn apart with a force $(7eF)V$, but at a separation d , significantly larger than the column diameters, they experience a mutually attractive force $2(7eF)^2/\epsilon ld$, where l is the length of the columns and ϵ is the dielectric constant of the liquid. Thus for the applied field to succeed in drawing the charges apart

$$7eFV > 2(7eF)^2/\epsilon ld,$$

or

$$d > (14e/\epsilon l) \cdot F/V = 2.8 \times 10^{-4} \text{ cm.}$$

The charge clouds were considered to have diameters less than one-tenth of this value, so that the applied field only competes favourably with the mutually attractive force when the charge clouds are already well separated. This indicates that the observed charge collection cannot occur by the Jaffé-Kramers process.

In the theory considered here the kinetic energy imparted to the electrons by the ionization process plays an important part in overcoming their strong attraction to the positive ion column. The fraction of the α -particle energy which goes into kinetic energy of the electrons is estimated in Section IV. The electrons lose their kinetic energy only after many collisions, of the order 10^5 in argon (a discussion of this point is given by Williams (1957)) so that they form a diffuse cloud of greater diameter than the positive ion column.

The electron cloud remains in dynamic equilibrium in the central field of the positive charge, while it is gradually reduced by recombination in the central region and loss of energy by collision of electrons with atoms.

In an external field perpendicular to the column of ionization the centres of the positive and negative charge clouds are separated or polarized by an amount which is calculated in Section III. This enhances the probability of escape from the column of the more energetic electrons, which during their random motion have sufficient energy to reach the critical radius at which the applied field becomes greater than the central field.

The motion of the electrons is such that, if a particular electron has sufficient energy, it will escape from the column merely because by chance it will eventually cross the critical radius in a direction favouring escape. The remaining less energetic electrons are more tightly held by the excess positive charge of the column, so that, however long the column lasts before recombination destroys it, it can make no further contribution to the ionization current but drifts bodily to the negative electrode.

A rigorous analysis of the model presented is virtually impossible since it depends essentially upon the initial distribution of electron energies, itself only

obtainable by numerical methods (e.g. Erskine 1954). This distribution is further complicated by the fact that the fastest electrons cause secondary or feather ionization, in tracks of low density, thus producing slow electrons well clear of the main column of ionization.

However, the main features of the theory are examined quantitatively in the following sections.

II. STRUCTURE OF AN ION COLUMN

The mean free path of positive ions is short, and collisions of heavy ions with neutral atoms rapidly reduce the kinetic energy of the ions to thermal energy. With respect to electrons, which are slowed down only after many collisions, the positive ions are thus considered to remain in a stationary column. The kinetic energy of the electrons then enables them to expand to a cloud of larger diameter against the potential energy of attraction to the positive ions. The density distribution of electrons in the ion column is not needed for approximate calculations, but the potential energy of a simple distribution can be calculated as a demonstration of the magnitude of the electron cloud expansion which is possible with a moderate electron kinetic energy.

The field F , at a distance r from the centre of any cylindrically symmetrical charge distribution of infinite length is due to the charge per unit length, ne/l inside a radius r , and has the value

$$F = 2ne/\epsilon lr, \dots\dots\dots (1)$$

where ϵ is the dielectric constant of the medium. Charge outside r does not contribute to the field at r . An electron in the ion column which is "expanding" past the radius r experiences a retarding field according to equation (1), where (ne) is the net positive charge inside r and l is the length of the column, supposing it to be much greater than the diameter. If r is outside the positive ion column, then n is also the number of electrons which have already expanded past r .

For the purpose of making an estimate of the potential energy of an "expanded" electron cloud for comparison with the kinetic energy available to the electrons, we may consider an ordered collapse of a uniformly dense cylinder, radius R , of N electrons into a smaller uniform cylinder radius R_0 of positive ions, exerting a field as in equation (1). The electrons inside R_0 may be considered as neutralizing (NR_0^2/R^2) positive ions, leaving $N(1 - R_0^2/R^2)$ positive ions inside R_0 and the same number of electrons outside. The collapse occurs by electrons crossing R_0 , beginning with the innermost shell, and neutralizing the first available positive ions.

An elementary cylinder of electrons, thickness du , collapsing from a radius u , experiences a central field

$$F = \frac{2Ne(1 - u^2/R^2)}{\epsilon lr} \text{ at radius } r.$$

It reaches the first un-neutralized positive ions at radius v given by

$$u^2/R^2 = 1 + v^2(1/R^2 - 1/R_0^2).$$

The number of electrons starting from the element du is $N \cdot 2u du/R^2$, so that the energy of collapse of electrons from this element is

$$dE = \frac{4N^2 e^2 u (1 - u^2/R^2) du}{\epsilon l R^2} \int_v^u \frac{dr}{r} = \frac{4N^2 e^2}{\epsilon l R^2} u \left(1 - \frac{u^2}{R^2}\right) \ln \left(\frac{u(1/R_0^2 - 1/R^2)^{\frac{1}{2}}}{(1 - u^2/R^2)^{\frac{1}{2}}} \right) du.$$

The total energy of collapse is found by integrating with respect to u between R_0 and R :

$$E = (N^2 e^2 / \epsilon l) \{ \ln (R/R_0) - \frac{1}{2} + \frac{1}{2} (R_0/R)^2 \}. \quad \dots\dots\dots (2)$$

Numerical solution of equation (2) shows that if the electrons reach a total potential energy of 0.5 MeV, they could expand to a uniform cylinder of radius R approximately three times the radius R_0 of the positive ion column.

This result can serve only as an order of magnitude because it is not imagined that uniform distributions describe the actual state of the ion columns.

III. POLARIZATION OF THE ION COLUMN

If the electron density in the central region is uniform the polarization or separation of the centres of the positive ion and electron distributions by an applied field is readily calculated. The density of electrons is defined by the characteristic radius R , the radius of the cylinder which the electrons would occupy if the whole distribution were of the same density as the central region.

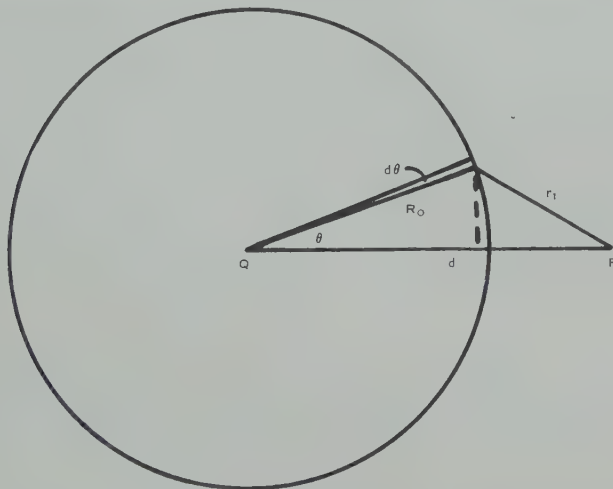


Fig. 1.—Calculation of the attractive force between overlapping cylinders of positive and negative charge.

Consider the force of attraction of an elementary hollow cylinder of dN positive ions, with radius R_0 and centre Q , to the centre P of a uniform cylindrical cloud of electrons, at a distance d from Q as in Figure 1. The component in the direction PQ of the force on an element $d\theta$ at θ is

$$df = \frac{2N(r_1^2/R^2)e \cdot dNe(d\theta/2\pi)}{\epsilon l r_1} \cos \phi,$$

since only the electrons inside radius r_1 are effective, where

$$r_1 \cos \varphi = d - R_0 \cos \theta.$$

Integrating over θ :

$$f = \frac{2Ne^2d \cdot dN}{\epsilon l R^2}.$$

Since this is independent of R_0 , the force on a solid cylinder of N positive ions with any radial distribution is $2N^2e^2d/\epsilon l R^2$.

This is equated to the force of the applied field, F , tending to pull the electrons away from the positive ions. Assuming n of the N electrons to have escaped from the column, this is

$$\begin{aligned} F(N-n)e &= 2N^2e^2d/\epsilon l R^2, \\ d &= \epsilon l R^2 F(N-n)/2N^2e. \end{aligned} \quad \dots\dots\dots (3)$$

The parameter R is a function of n , since, by virtue of the escape of n electrons, the density in the central region is reduced. However, the fastest electrons would have spent only a small fraction of the time in the central region and not very greatly added to the density; also the net positive charge remaining tends to reduce the size of the electron cloud slightly, so that the overall error which results from taking R as constant is not great.

IV. ENERGY SPECTRUM OF THE ELECTRONS

The process of expansion of the electron cloud is similar to diffusion; each electron may traverse the positive ion column several times before the equilibrium state is reached. At equilibrium each electron conserves its total (kinetic plus potential) energy, the potential energy being gained by motion outwards against the central field described by equation (1), in which n is a function of r , determined by the electron distribution. During the expansion this distribution is time-dependent, so that any attempt to derive it from a given initial distribution of electron kinetic energies would be difficult, and this would be a necessary preliminary to determining the loss of kinetic energy during expansion and the ionization current. Here a much simpler calculation of the electron energy spectrum from measured ionization currents is made.

If the n most energetic electrons escape from the column in a field F perpendicular to the column, then the remaining cloud of $(N-n)$ electrons has a critical outer radius r_c at which the total field is zero in the direction of F . The $(n+1)$ th most energetic electron does not quite have enough initial kinetic energy to reach r_c .

Allowing for polarization as in Section III, we have for the field at r_c

$$0 = \frac{2Ne}{\epsilon l(r_c + d)} - \frac{2(N-n)e}{\epsilon l r_c} - F,$$

where the first term is attraction to the N positive ions, the second term is repulsion by the $(N-n)$ remaining electrons, and the third term is the applied field. This gives

$$r_c = \left[\frac{ne}{\epsilon l F} - \frac{\epsilon l F(N-n)}{4N^2e} R^2 \right] + \sqrt{\left\{ \left[\right]^2 - \frac{(N-n)^2}{N^2} R^2 \right\}}. \quad \dots\dots (4)$$

A maximum value of R is set by the conditions that r_c must be real at all values of F . Measurements at the highest field used by Williams and Stacey (1957) give

$$R < 0.9 \times 10^{-5} \text{ cm.}$$

In the following numerical calculations the value $R = 0.75 \times 10^{-5} \text{ cm}$ is taken; this is negligible at all but the highest fields and indicates that polarization is only important at high fields.

TABLE 1
CALCULATION OF SPECTRUM OF ELECTRON ENERGIES

F (V/cm)	n	r_c (10^{-5} cm)	ΔE (eV)	$E - E'$ (eV)	$-\frac{1}{N} \frac{dn}{dE}$
1	4.90	19.3	0.494×10^{-3}	5.29_{48}	0.139
3.162	18.6	23.3	1.41×10^{-3}	5.29_{43}	0.139
10	57.6	22.8	5.29×10^{-3}	5.29_{29}	0.138
31.62	2.04×10^2	25.5	2.21×10^{-2}	5.28_{76}	0.139
100	7.25×10^2	28.6	4.86×10^{-2}	5.26_{55}	0.109
316.2	1.70×10^3	21.3	0.126	5.21_{69}	0.095
10^3	3.99×10^3	15.7	0.258	5.09_{06}	0.0881
3.162×10^3	8.32×10^3	10.4	0.485	4.83_{24}	0.0812
10^4	1.68×10^4	6.56	0.955	4.34_{73}	0.0751
3.162×10^4	3.02×10^4	3.65	1.64	3.39_2	0.0670
10^5	5.31×10^4	1.89	1.75	1.75_3	0.0721
3.162×10^5	9.12×10^4	0.815		0	0.1772

If n electrons reach r_c in a field F and $(n + dn)$ electrons reach $(r_c + dr_c)$ in a field $(F + dF)$ then the energy difference dE between the n th and $(n + dn)$ th most energetic electrons is

$$dE = (2ne^2/\epsilon l r_c) dr_c - \{(F + dF)(r_c + dr_c - r_0)e - F(r_c - r_0)e\},$$

$$dE/dn = (2ne^2/\epsilon l r_c - eF) dr_c/dn - (r_c - r_0) e dF/dn, \quad \dots \dots \dots (5)$$

where r_0 is the effective radius from which the electrons have expanded, for which the r.m.s. radius of the positive ion column is taken. To obtain an approximate value of r_0 we may put $r_0 = R_0/\sqrt{2}$, where R_0 is the characteristic radius of the positive ion column. Adopting the result of Section II, $R_0 = \frac{1}{3}R$, we have

$$r_0 = R/3\sqrt{2} \simeq 0.2 \times 10^{-5} \text{ cm.}$$

In equation (5) n is the number of electrons with energy E or greater, equated to the number of electrons escaping to give the observed ionization current in an applied field F .

Values of r_c were calculated from equation (4), using the results of Williams and Stacey (1957). These were used to find dr_c/dn graphically and hence dE/dn , from which the spectrum of electron energies shown in Figure 2 was obtained by numerical integration. Details of the calculation are given in Table 1. E' represents an arbitrary energy because the lower end of the spectrum is experimentally inaccessible, but a reasonable estimate of E' can be made from the fact that the area under the curve must integrate to unity.

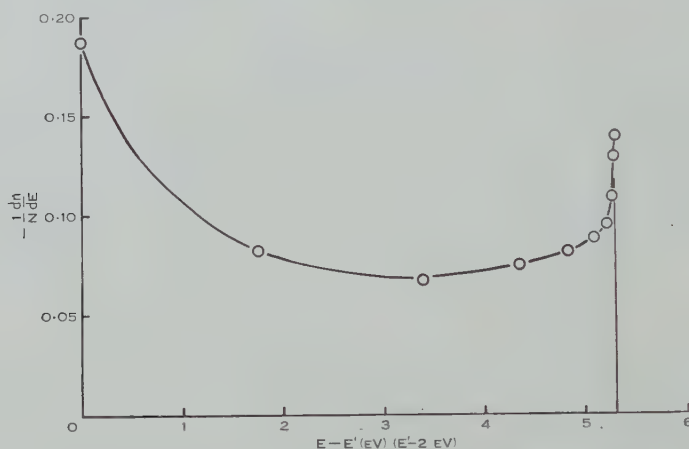


Fig. 2.—Energy spectrum of electrons in α -ionized liquid argon.

Drawing the spectrum with a sharp cut-off at $(E - E') = 5.3$ eV is necessitated by the fact that the experiments extended down to the collection of 5 electrons per α -particle, so that the area under the curve beyond the last experimental point is almost invisibly small.

The peak extending from 5.3 eV downwards is evidently a result of secondary ionization, by which the very energetic electrons produce secondaries well clear of the main column of ionization. In terms of the present theory the secondary electrons are produced with appreciable potential energy. The spectrum of Figure 2 thus represents final energies of electrons in the column, being a modification of the spectrum of initial kinetic energies. The sharpness of the cut-off at $(E - E') = 5.3$ eV is evidently due to an excitation of argon atoms at this energy, so that more energetic electrons rapidly lose their excess energy.

The cut-off is clearly at a lower energy than the first ionization energy of argon at 16 eV. This would require the curve of Figure 2 to double back to low values of $(1/N)dN/dE$ at low E in order that the area under the curve shall integrate to unity, and then the total energy represented by the spectrum would be at least 2 MeV. Since this would allow no energy for excitations, double ionization, and positive ion kinetic energy, it is clearly inadmissible, although the results of Ward (1954) indicate that the energy lost by excitation is small.

More reasonable is the simple extrapolation of the curve to the point at which the area under it is unity. This indicates $E' \sim 2$ eV, in which case the total energy in the spectrum is approximately 0.5 MeV. This is the result used in Section II.

V. DISCUSSION

In the preceding sections a number of assumptions have been made, requiring more detailed consideration.

Ionic recombination has been ignored, but its importance depends upon the time scale of the processes described. However, the mean free path for scattering of the electrons can be calculated from the electron mobility data of Williams (1957). It varies with energy, being of the order 10^{-5} cm, so that an electron would be expected to experience an average of one scattering collision per traversal of the positive ion column. The density of positive ions in the column averages less than 10^{-5} of the neutral atom density, so that encounters with positive ions are of the order 10^{-5} times as frequent. The electrons would therefore need to make some 10^4 traversals of the positive ion column during the expansion for recombination to have a noticeable effect on ionization current, and the relatively long mean free path clearly indicates that this is highly improbable.

That the electron cloud should retain a cylindrical form in an applied field is a simplifying assumption. The cloud will be misshapen but the polarization considered in Section III gives an approximation to the field experienced by an escaping electron. The fact that the polarization effect is negligible except in high fields suggests that only at high fields (low electron energies) could the results of the present calculation be in error on account of distortion of the electron cloud.

In deriving equation (5) it was assumed that the energy expended by an electron in moving outwards against the central field of the positive ions is independent of the applied field. The correctness of this assumption depends upon the mechanism of electron cloud expansion. At low fields an electron of average energy may be partially screened from the positive ions by more energetic electrons, which are temporarily at a smaller radius, but which would escape from the column at an early stage in the expansion in a high field. The magnitude of this effect is not known, so that in equation (5) it is assumed that the expansion of electrons to any particular potential energy has the same form independent of the applied field.

Experimental data are not available for ionization currents due to α -particles from collimated sources. These would be necessary for correct application to the theory presented which concerns only the escape of electrons in fields perpendicular to the α -particle tracks. Jaffé (1913) recognized this problem, but merely stated that the fields considered in his theory were the components of the applied fields perpendicular to the particle tracks, i.e. that escape from the ends of the columns of ionization could be neglected. This was reasonable in the Jaffé-Kramers theories, since the probability of recombination during the drift of an ion through a large fraction of the ions of opposite sign was almost

unity. In the present theory, recombination is much less probable since the positive ions and electrons have very little overlap and therefore escape of electrons in a field parallel to the column cannot be neglected altogether. The pulse experiments of Williams and Stacey (1957) indicate that the direction of α -particle emission with respect to the applied field does influence ionic collection, although the effect is smaller than that of Jaffé's theory. This effect is not considered in the foregoing theory, which therefore uses a "smoothed out" ionization current curve. Its influence on the calculations may introduce an error up to 50 per cent. in the estimated radii of the column and the energy spectrum.

The estimate of r_0 which was made in Section IV must be regarded as a lower limit, since it was obtained by considering the characteristic radius R of the central region of the electron cloud to enclose all the electrons, whereas it merely reflects the density of the central region. The upper limit is set by $r_0 = R = 0.75 \times 10^{-5}$ cm. This range of values puts an additional 20 per cent. possible error on to the value of $-(1/N)dn/dE$ at $(E-E')=0$ in Figure 2, in which the upper limit is indicated.

The values of n and r_c in Table 1 do not give the electron density at any radius in zero field, since r_c has been calculated as the radius which n electrons can reach with the assistance of an external field. In the absence of an external field, the extreme radius of the column is approximately 20×10^{-5} cm.

A primary electron energy spectrum may be inferred from the results of calculations by Erskine (1954), who plotted the cross section Q for ionization of helium atoms as a function of the wave number k of the liberated electrons. The distribution may be approximately represented by a simple analytical expression

$$Q = Ck \exp(-k^2/\sigma^2),$$

where C and σ are constants. This leads directly to a spectrum of primary electron energies:

$$\frac{dN}{N} = \frac{1}{E_0} \exp(-E/E_0) dE,$$

in which E_0 is the average energy. When modified by a cut-off energy and secondary maximum as discussed in Section IV, this is not inconsistent with the appearance of Figure 1, but the energy range indicated by Erskine (wave numbers of the order 10^8 cm^{-1}) is clearly too large, giving an average electron energy in excess of 100 eV.

Liquids or gases with positive electron affinities evidently still require the Jaffé-Kramers treatment, although a modification of the assumption of identical positive and negative ion distributions must be strongly suggested. A large diameter of the negative ion cloud would result if the electrons began the expansion process as in argon before becoming attached to neutral atoms. This allows a larger ionization current than the original Jaffé-Kramers theories and avoids the objection to these theories which is raised in Section I.

VI. REFERENCES

- ERSKINE, G. A. (1954).—*Proc. Roy. Soc. A* **224**: 362.
GERRITSEN, A. N. (1945).—*Physica* **14**: 381, 407.
JAFFÉ, G. (1913).—*Ann. Phys., Lpz.* **42**: 303.
KRAMERS, H. A. (1952).—*Physica* **18**: 665.
WARD, A. (1954).—*Proc. Phys. Soc. Lond. A* **67**: 841.
WILLIAMS, R. L. (1957).—*Canad. J. Phys.* **35**: 134.
WILLIAMS, R. L., and STACEY, F. D. (1957).—*Canad. J. Phys.* **35**: 928.

THERMAL STRUCTURES IN THE LOWEST LAYERS OF THE ATMOSPHERE

By R. J. TAYLOR*

[*Manuscript received January 22, 1958*]

Summary

An analysis has been made of a number of records of simultaneous fluctuations of temperature at four heights. The results show that the correlation between temperatures at two heights is greatest when the records are displaced relatively to each other, so that earlier points in the upper record are considered together with later points in the lower, by an amount depending on the mean wind speed difference between the two heights. The correlation increases with height relatively to what would be expected in a locally isotropic turbulence.

This and other evidence points to the existence of organized thermal structures of considerable vertical extent which are superimposed on a background of random turbulent disturbances. It is probable that these structures represent convection plumes.

I. INTRODUCTION

During the summer and autumn of 1956 a number of simultaneous records of temperature at four different heights were made in order to investigate a phenomenon first noticed by the author in 1949 and briefly mentioned by Priestley (1952). It was found then that, when the correlation coefficient is calculated between simultaneous records of temperature at two heights, the maximum value is attained, not when the correlation is between temperatures at corresponding instants, but when earlier temperatures at the upper level are taken with slightly later temperatures at the lower. The general impression given is that temperature anomalies which are big enough to affect both recording points in general reach the higher levels first. A similar effect for velocity fluctuations in a wind-tunnel boundary layer has been reported by Favre, Gaviglio, and Dumas (1957).

The earlier observations mentioned were confined to two simultaneous temperature records (one thermometer fixed at either 1 or 2 m above the ground and the other at a variable height) and were not accompanied by detailed wind and temperature profile measurements. It was therefore decided to make a completely new set of observations with temperature fluctuations at four heights and vertical velocity fluctuations at one of them, in order to permit a qualitative comparison of the natures of the two types of record.

II. RECORDING AND ANALYSIS

Temperatures were measured using fine wire resistance thermometers of approximately $2000\ \Omega$ of 0.0005 in. diameter nickel wire included in suitable

* Division of Meteorological Physics, C.S.I.R.O., Aspendale, Vic.

Wheatstone bridge networks provided with zero-shift controls and fixed resistors for rapid calibration-checking. The maximum dimension of each thermometer coil was less than 2 in. Since temperature fluctuations only were of interest, the thermometer bridges were calibrated for sensitivity but not for absolute value. The anemometer used for measuring the vertical velocity fluctuations was of a new design which has been described elsewhere (Taylor 1958). The records were made using the galvanometers and photographic arrangements incorporated in the atmospheric fine structure recorder described by McIlroy (1955). The thermometers were mounted at 1.5, 4, 16, and 30 m above the ground and the anemometer at 1.5 m. Response times of all the instruments were so small that the resolution obtained in the traces is determined by the recording galvanometers (natural period 2 sec) rather than by the sensing parts. All records were made over 10-min periods.

Simultaneously with the fluctuation recordings, measurements were made of mean wind speed at a number (eight or nine) of heights and of mean temperature difference over four height intervals.

The records were in the form of sheets of photographic paper about 45 cm by 200 m, each carrying traces of temperature at the four heights ($T_{1.5}$ etc.) and of vertical component at 1.5 m (w). The analysis was carried out on the differential analyser described by Taylor and Webb (1955), the quantities calculated being:

- (1) Root-mean-square fluctuations denoted by $\sigma_{T(1.5)}$ etc.
- (2) Correlation coefficients R between temperatures at two heights as a function of time-lag (τ) between the two traces. Special interest attaches to three values; R at zero lag (R_0), the maximum correlation (R_m), and the lag τ_m when $R=R_m$. Values of τ are counted positive when earlier times at a higher level are coupled with later times at a lower.
- (3) On four occasions, a graph of the progressive total of the covariance between temperatures at two heights was written by the analyser as a function of record length x . That is, the relationship of $\int_0^x T'_{1.5} T'_4 dx$ (and similar quantities) with x was displayed, where the primes indicate deviations from means. This permitted the location on the original record of those features contributing most strongly to the correlations.

Each record analysed (with certain exceptions) thus gave five values of σ and six sets of values of R at various τ . Because of the time-consuming nature of the calculations only 20 complete records (16 day-time observations and 4 from the evening) were analysed in this way. The results, however, are sufficiently consistent for valid conclusions to be drawn.

III. RESULTS

In almost every case the values of R showed a clear-cut maximum at a moderate positive value of τ (mostly between 1 and 10 sec). Qualitative inspections of the day-time records clearly show this time lag at maximum

correlation, and obvious "structures" affecting all heights not quite simultaneously are frequent. These structures appear on the records as asymmetrical triangular waves of temperature (gradual rise followed by sudden drop) and frequently attain several Centigrade degrees in magnitude with duration mostly about 10–20 sec. An example has been used by Priestley (1956) and the reader may be referred to his Figure 3 for illustration. An interesting feature of the temperature traces is that there frequently seem to be a fairly constant "base" temperature and temperature fluctuations which are almost exclusively positive relative to it. The effect is particularly well shown in a record obtained by I. C. McIlroy (unpublished data) in another connexion. Part of this record (which was made with $\frac{1}{2}$ sec galvanometers and refers to a height of 7 m above the ground) is reproduced in Figure 1. Although the asymmetrical, triangular nature of the temperature fluctuations is less evident here, the separation of temperature into quiet and disturbed periods is well marked, as is also the virtual absence of fluctuations negative with respect to the "quiet" temperature. Figure 1 also

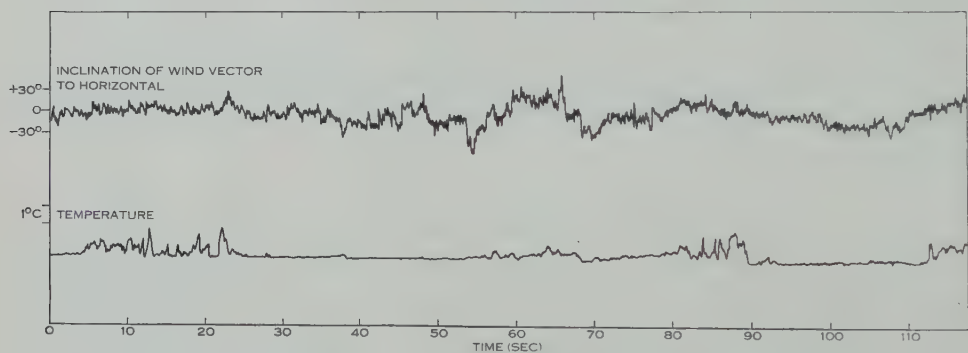


Fig. 1.—Record showing "quiet" and "disturbed" periods in temperature at 7 m.

shows the angle of inclination of the wind vector to the horizontal, and it is clear that the motion is predominantly upwards during the "disturbed" temperature periods although the inclination record is strongly influenced by random motion at all times. These features also appear in the w -traces of the present records.

The association between thermal structures and vertical velocity component is also very well shown in a record made at 23 m above the ground by E. L. Deacon (unpublished data), a part of which is reproduced here in Figure 2. Particularly striking is the virtual absence of negative contributions to the turbulent heat flux.

Of the 16 day-time observations (involving 96 (R, τ) graphs) only five negative τ_m were found and for three pairs of traces (all in the same record) it was impossible to locate a maximum R in any reasonable range of τ . The four evening runs analysed lacked a 4 m temperature record but showed positive τ_m in all but one case. In general, the evening runs differed from the day-time ones in the absence of structures such as those referred to above, in having less well-marked maxima, and in having an occasional very large τ_m . The last

two points are illustrated by calculating, separately for day-time and evening observations, the ratio of the average of $(R_m - R_0)$ to the square of the average τ_m . The values are 0.005 sec^{-2} (day-time) and 0.001 sec^{-2} (evening). Table 1 shows values of R_0 , R_m , τ_m , σ_T , and σ_w for those traces which have been analysed, together with the Richardson number at 1.5 m ($Ri_{1.5}$).

An obvious explanation for the existence of a positive τ_m lies in the distorting effect of the mean wind shear on any atmospheric structure having sufficient vertical extent. If there is any process in the flow creating temperature disturbances over the height range of the observations, then the wind shear will

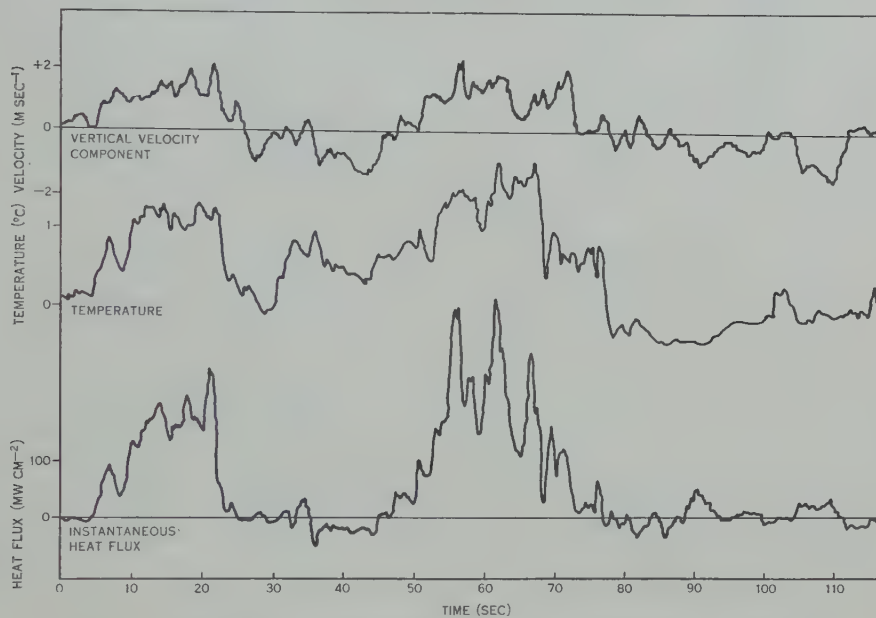


Fig. 2.—Association of temperature fluctuations with vertical velocity.

tend to give them a rotation in the observed sense. However, the narrowness of the range of τ_m observed makes it appear improbable that we are here dealing with the effects of wind shear on disturbances originally oriented at random in all possible directions.

The suggestion that wind shear is responsible for the observed displacement of maximum R is given weight by a detailed comparison of the profiles. The means of day-time observations of both τ_m and wind speed difference (Δu) show fairly smooth curves when divided by the height interval (Δz) concerned and plotted against the geometric mean of the extremities of that interval. Figure 3 shows these values of $\tau_m/\Delta z$ and $\Delta u/\Delta z$ expressed as fractions of the values over the whole height range 1.5–30 m. There is good agreement except in the layer closest to the ground where disturbing influences are generally likely to be greatest. The suggestion that the observed thermal structures are convection columns will be examined in more detail later; here it is sufficient to note that the general pattern of τ_m observed is consistent with the existence of convection

columns travelling down wind at a speed less than the wind speed at any of the heights of observations.

The mean values of correlation coefficients at zero lag from Table 1 are repeated in Table 2 which also shows the standard errors of the means as derived from the observed scatter in R_0 .

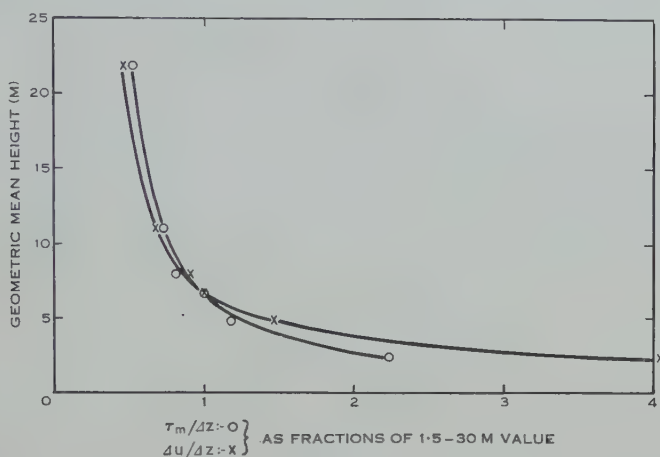


Fig. 3.—Profiles of τ_m and wind speed difference.

It will be noticed that R_0 is generally smaller by night than by day and, at least in the layer 16–30 m, there is no doubt that the difference is significant.

When considering the variations of R_0 with height, one must take account of the fact that the correlations are measured over unequal height intervals and some system of normalization must be adopted. The present author (Taylor 1955) has shown that, as far as the velocity field is concerned, the predictions of similarity theory apply up to eddy sizes which may be several

TABLE 2
MEANS AND STANDARD ERRORS OF CORRELATION COEFFICIENTS

Layer (m)	1.5-4	1.5-16	1.5-30	4-16	4-30	16-30
Day-time mean R_0 ..	0.68	0.34	0.29	0.48	0.38	0.65
Standard error of mean	0.02	0.04	0.04	0.04	0.03	0.02
Evening mean R_0 ..		0.24	0.20			0.31
Standard error of mean		0.03	0.03			0.08

times the height of observation. It is probably reasonable, therefore, to apply similar considerations to the analysis of the temperature field. Theoretical discussions of temperature fluctuations in locally isotropic turbulence have been made by Obukhov (1949) and Corrsin (1951). The former writes in terms of “structural functions”—the mean square temperature difference between two points—while the latter uses spectrum functions, but both agree in determining the temperature field over the intermediate range of wave numbers in terms of the distance involved (or wave number), the rate of dissipation of

kinetic energy (ε), and a quantity N which describes the rate at which temperature differences are being destroyed by the action of molecular conduction. N has dimensions (temperature)² × (time)⁻¹ and is defined by

$$N = k \overline{(\text{grad } T)^2},$$

where k is the thermal diffusivity and T is temperature. Dimensional analysis then indicates that

$$\overline{(T_1 - T_2)^2} = CN\varepsilon^{-\frac{1}{3}}r^{\frac{2}{3}},$$

where T_1 and T_2 are temperatures at two points separated by a distance r and C is a constant. Tatarsky (1956) discusses this expression, produces observations to demonstrate that $\overline{(T_1 - T_2)^2}$ does indeed vary as the two-thirds power of r , and relates the constant C to the mean meteorological conditions. In locally isotropic turbulence, the mean square temperature difference can be written in the form

$$\overline{(T_1 - T_2)^2} = 2\overline{T'^2}(1 - R),$$

where $\overline{T'^2}$ is the mean square temperature fluctuation and we therefore have

$$(1 - R) \propto r^{\frac{2}{3}}.$$

This relationship has been taken as the basis of a scheme of normalization for R_0 to make allowance for the varying height interval Δz , and Table 3 shows the means and standard errors of the means of $y = (1 - R_0)/\Delta z^{\frac{2}{3}}$.

TABLE 3
MEANS AND STANDARD ERRORS OF $y = (1 - R_0)/\Delta z^{\frac{2}{3}}$

Layer (m)	1.5-4	1.5-16	1.5-30	4-16	4-30	16-30
Day-time mean y (m ^{-$\frac{2}{3}$)}	0.174	0.112	0.076	0.100	0.071	0.060
Standard error of mean (m ^{-$\frac{2}{3}$)}	0.013	0.007	0.004	0.007	0.004	0.004
Evening mean y (m ^{-$\frac{2}{3}$)}		0.128	0.085			0.118
Standard error of mean (m ^{-$\frac{2}{3}$)}		0.006	0.003			0.013

It is clear from these results that, by day, y does decrease significantly with increasing height in a systematic way and therefore that the correlations are relatively greater with height than would be expected in a locally isotropic turbulence. The presence of some sort of eddy, directionally influenced by the boundary, is thus indicated.

By night, the picture is less straightforward. There is some evidence for a significant difference between the layers 1.5-16 m and 1.5-30 m, but the layer 16-30 m provides no support for the trend thus suggested. The difference between day and night in the 16-30 m layer is very clear.

The suggestion that the directional influence of the boundary is greater by day than by night is consistent, of course, with the fact that the vertical turbulent fluxes of heat and momentum are larger in the former case.

An example of a graph of progressive total of covariance is shown in Figure 4, which gives $\int_0^x T'_{1.5} T'_4 dx$ as a function of record length x for Run 1 of January 17, 1956 with $\tau = \tau_m$. Parts of the curve are obviously much steeper than the general trend and these parts are those which contribute particularly strongly to the

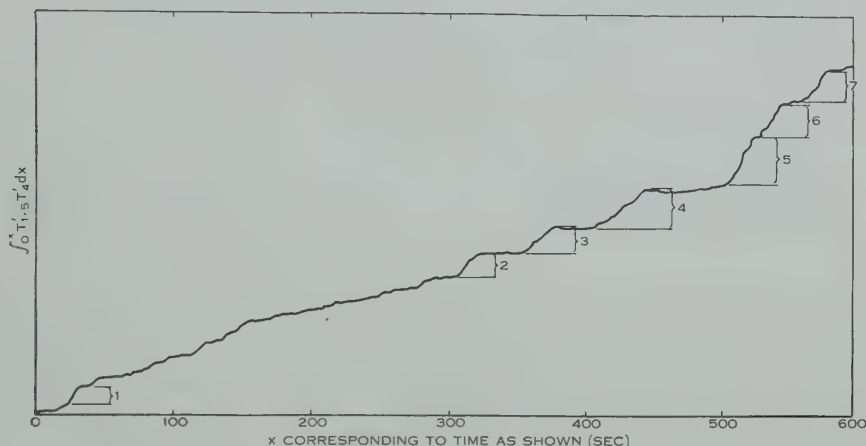


Fig. 4.—Events contributing most strongly to covariance.

covariance. There is clearly some subjectivity in defining them but seven "events" have been chosen (as indicated by braces and numbers in Figure 4) and, for this run, these events contribute 63 per cent. of the covariance in 24 per cent. of the total recording time. Table 4 summarizes the results from the four records so analysed.

TABLE 4
EVENTS CONTRIBUTING STRONGLY TO COVARIANCE

Date	Run No.	Height Interval (m)	Fraction of Covariance Contributed p	Fraction of Time Occupied q	$\frac{q}{p}$	No. of Events
17.i.56	1	1.5-4	0.632	0.242	0.383	7
17.i.56	2	1.5-4	0.652	0.223	0.342	9
17.i.56	2	1.5-16	0.734	0.191	0.260	8
17.i.56	2	16-30	0.364	0.089	0.244	4

An examination of the graph of cumulative covariance in conjunction with the original records discloses that almost all the special events listed in Table 4 can be identified with structures similar to those referred to above. The few cases where this is not so are associated with broader oscillations of temperature of rather smoother character. As far as they go, the results shown in Table 4 would indicate that these special events become rarer, but more intense, with increasing height.

IV. DISCUSSION

The evidence provides a convincing picture of a turbulent flow of the usual random character, on which are superimposed organized structures of considerable vertical extent. Briefly recapitulating, the main points in this evidence are :

- (i) the fact that temperature excursions affecting all four heights nearly simultaneously can be identified on the original records,
- (ii) the presence on the graphs of cumulative covariance of special events corresponding, in most cases, to these excursions,
- (iii) the increase with height of correlation coefficient relative to what would be expected in a locally isotropic turbulence,
- (iv) the fact that the temperature excursions affect all heights at times which are consistent with their being continuous vertically but distorted by wind shear.

The observations, however, let us go further than this. The structures do not appear by night and when they do appear, by day, they are generally associated with upward vertical velocity component at all heights. Moreover, there exist occasions (similar to that illustrated in Figure 1) when these structures are almost the only temperature disturbances existing and then they clearly take the form of positive deviations from a nearly constant base temperature. These considerations make it extremely probable that they represent some sort of convection process. Priestley (1956) has discussed the various sorts of possible convection plume and the present observations are generally consistent with the models he there proposes.

V. ACKNOWLEDGMENTS

The original observations were taken with the assistance of many members of the staff of the Division of Meteorological Physics and the very laborious computations were carried out under the direction of Mr. N. E. Bacon.

VI. REFERENCES

- CORRSIN, S. (1951).—*J. Appl. Phys.* **22** : 469.
 FAVRE, A. J., GAVIGLIO, J. J., and DUMAS, R. (1957).—*J. Fluid Mech.* **2** : 313.
 MCILROY, I. C. (1955).—C.S.I.R.O. Aust. Div. Met. Phys. Tech. Pap. No. 3.
 OBUKHOV, A. M. (1949).—Bull. Acad. Sci. U.R.S.S. (Geog. & Geophys. Ser.) **13** : 58.
 PRIESTLEY, C. H. B. (1952).—Geophys. Res. Pap., M.I.T. No. 19, p. 33.
 PRIESTLEY, C. H. B. (1956).—*Proc. Roy. Soc. A* **238** : 287.
 TATARSKY, V. I. (1956).—Bull. Acad. Sci. U.R.S.S. (Geophys. Ser.) No. 6, p. 689.
 TAYLOR, R. J. (1955).—*Aust. J. Phys.* **8** : 535.
 TAYLOR, R. J. (1958).—*J. Sci. Instrum.* **35** : 47.
 TAYLOR, R. J., and WEBB, E. K. (1955).—C.S.I.R.O. Aust. Div. Met. Phys. Tech. Pap. No. 6.

THE WILSON EFFECT IN SUNSPOTS

By R. E. LOUGHHEAD* and R. J. BRAY*

[*Manuscript received December 20, 1957*]

Summary

New observations of the Wilson effect in a small regular sunspot have demonstrated the reality of the phenomenon. The measurements are based on 38 photographs selected from some 24,000 obtained during the passage of the spot from the east to the west limb.

It is pointed out that the changing appearance of a sunspot during its passage across the disk is essentially a problem in the theory of radiative transfer. On the basis of a simple model a qualitative explanation is found for the Wilson effect and for the anomalous foreshortening of sunspot areas.

I. INTRODUCTION

The question of the structure in depth of sunspots is of considerable interest both in connexion with the properties of the photosphere and the mode of formation of the spots themselves (Cowling 1953). The changing appearance of a spot during its passage across the disk must depend on the absorption and emission coefficients of the spot material and of the photosphere, and on the variation of these quantities with depth. In particular, observations of the foreshortening effect as the spot approaches the limb of the Sun may be expected to provide some elucidation of this structure.

Many solar observers have noticed that, as a spot approaches the limb, the width of the penumbra on the inner side decreases at a greater rate than that on the side of the spot nearer the limb (Abetti 1957, cf. p. 65). This phenomenon was first described by A. Wilson in 1769 and is termed the Wilson effect. The work of subsequent observers has in general confirmed the reality of the effect, although the observed magnitude has decreased with the application of more modern techniques. Discrepancies among the results of the early observers can be attributed mainly to the smallness of the quantities to be measured. Near the limb the width of the highly foreshortened penumbra may be reduced to 2 sec of arc or less; this can be measured only with the aid of a high-resolution instrument under conditions of good atmospheric seeing. In addition, by confining their measures to the limb, many early workers failed to take account of any possible asymmetry in the spot. To avoid systematic error due to this cause (cf. Abbot 1912) measurements should be made at both limbs and at the centre, where the foreshortening is negligible.

In recent times Unsöld (1955) and Waldmeier (1955, cf. p. 169) have even inclined to the view that the effect has a psychological origin.

* Division of Physics, C.S.I.R.O., University Grounds, Chippendale, N.S.W.

In the present paper observations are described of a small, fairly regular, stable sunspot which was photographed on all but one day of its passage across the disk with a 5 in. photoheliograph. The 38 photographs upon which the measurements are based were selected from some 24,000 obtained during this period. The interpretation of the observations is briefly considered in the final section.

II. OBSERVATIONS

The spot selected for this study was first seen at the east limb on May 19 and last seen at the west limb on May 30. During this period it was photographed repeatedly (except on May 27, when cloud intervened) with the 5 in. photoheliograph (Loughhead and Burgess 1958) of the C.S.I.R.O. Division of Physics Solar Observatory. This instrument is designed to photograph any selected region of the solar disk on 35 mm film at 5-sec intervals; the diameter of the solar image is 20 cm. The effective wavelength is 5400 Å with a bandwidth of 800 Å.

Plate 1 shows the changing appearance of the spot during its passage across the disk. Each photograph is orientated so that the direction of greatest foreshortening is horizontal. It will be noticed that, although the spot is fairly regular when seen near the central meridian, at both the east and west limbs the width of the penumbra on the side of the spot nearer the limb is significantly greater than that on the other side. This illustration leaves no doubt as to the reality of the Wilson effect for this particular spot; it remains to determine its magnitude.

III. REDUCTION

The quantities required are the apparent widths of the penumbra AN and $A'N'$ (Fig. 1) measured along the line of greatest foreshortening on the sides of the spot directed towards and away from the solar limb respectively. This line lies along the arc of the great circle whose plane contains the solar radius and the line-of-sight to the centre of the spot.

The ratio

$$f = AN/A'N', \dots\dots\dots (1)$$

which varies with the heliocentric angle θ of the spot, is a measure of the Wilson effect. A minor difficulty arises in that the spot, as seen from the Earth, appears to rotate about the direction of greatest foreshortening as it moves across the disk. Measurements made on successive days along this line therefore refer to different lines in the spot itself. However, for the 3 days during which the spot is close to the limb, the apparent rotation is so small ($\sim 2^\circ$) that it may safely be ignored. Measurements of f were first made for the 3 days during which the spot was close to the east limb, and the results compared with measurements for the two days during which the spot was near the central meridian, where the foreshortening is negligible. The line in the spot corresponding to the limb measurements was transferred to the centre photographs with the aid of Stonyhurst disks. All measurements were therefore made along the same line in the spot. A similar set of measurements was made for the west limb and the results also compared

with the centre observations. For the reasons mentioned above the east and west limb measurements were made along different lines in the spot, the angle between the lines being approximately 7° .

The 38 photographs upon which the measurements are based were selected from some 24,000 contained in the 35 mm records. This selection provided five good photographs for each day, except May 19, when only three of sufficient quality were obtained. By choosing from such a large number of photographs it

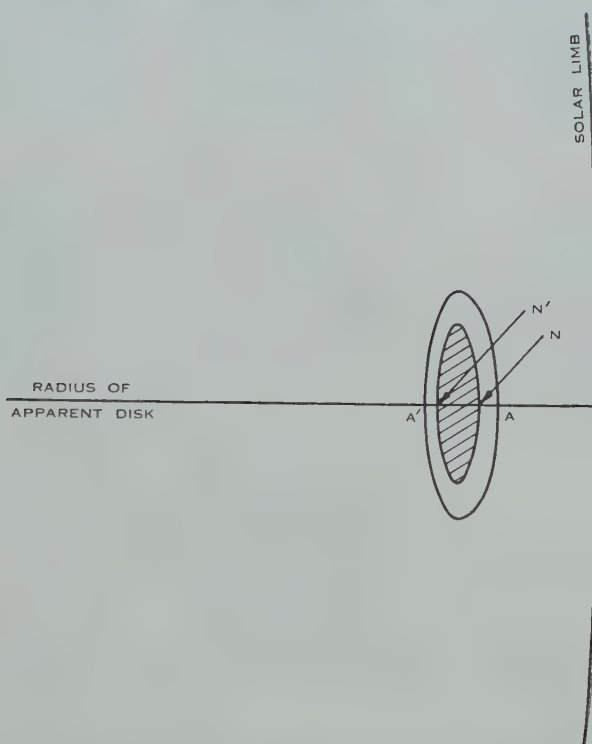


Fig. 1.— AA' is the arc intercepted by the spot on the great circle whose plane contains the solar radius and the line-of-sight to the centre of the spot. AN and $A'N'$ are the apparent widths of the penumbra on the sides of the spot towards and remote from the solar limb respectively.

is usually possible to obtain several of high quality even on days of relatively poor seeing. Enlargements were made of the 38 photographs on such a scale that 1 mm corresponded to 0.9 sec of arc. Kodak waterproof bromide paper was used to avoid possible distortion during processing.

The technique of measurement consisted of tracing the outlines of the umbra and penumbra from the enlargements onto transparent paper. In tracing the penumbra irregularities in the outer boundary were averaged out and similarly, in tracing the umbra, small dark projections into the penumbra were ignored. No difficulty was experienced except at the extreme limb where foreshortening reduces the apparent width of the penumbra to a few seconds of arc. In addition,

at the extreme limb, the boundary between the umbra and penumbra, while quite sharp on the side nearer the limb, appears rather diffuse on the other side (cf. Plate 1). This phenomenon has been noticed in other sunspots near the limb and merits further study. To avoid either over- or underestimating the width an attempt was made to trace upper and lower limits to the boundary, the mean then being taken. For the limb photographs the penumbral widths were measured along the line of greatest foreshortening. This line in the spot was

TABLE 1
PENUMBRAL MEASUREMENTS*

(a) East Limb

	East Limb			Centre	
Date θ	19.v.57 73.1°	20.v.57 59.3°	21.v.57 45.2°	24.v.57 8.6°	25.v.57 12.1°
AN	$4.1 \pm 0.2''$	$5.6 \pm 0.3''$	$6.4 \pm 0.3''$	$9.3 \pm 0.3''$	$7.9 \pm 0.6''$
$A'N'$	$2.2 \pm 0.2''$	$3.5 \pm 0.3''$	$5.4 \pm 0.2''$	$8.1 \pm 0.3''$	$8.6 \pm 0.3''$
f	1.88 ± 0.15	1.57 ± 0.15	1.18 ± 0.07	1.15 ± 0.06	0.91 ± 0.08

(b) West Limb

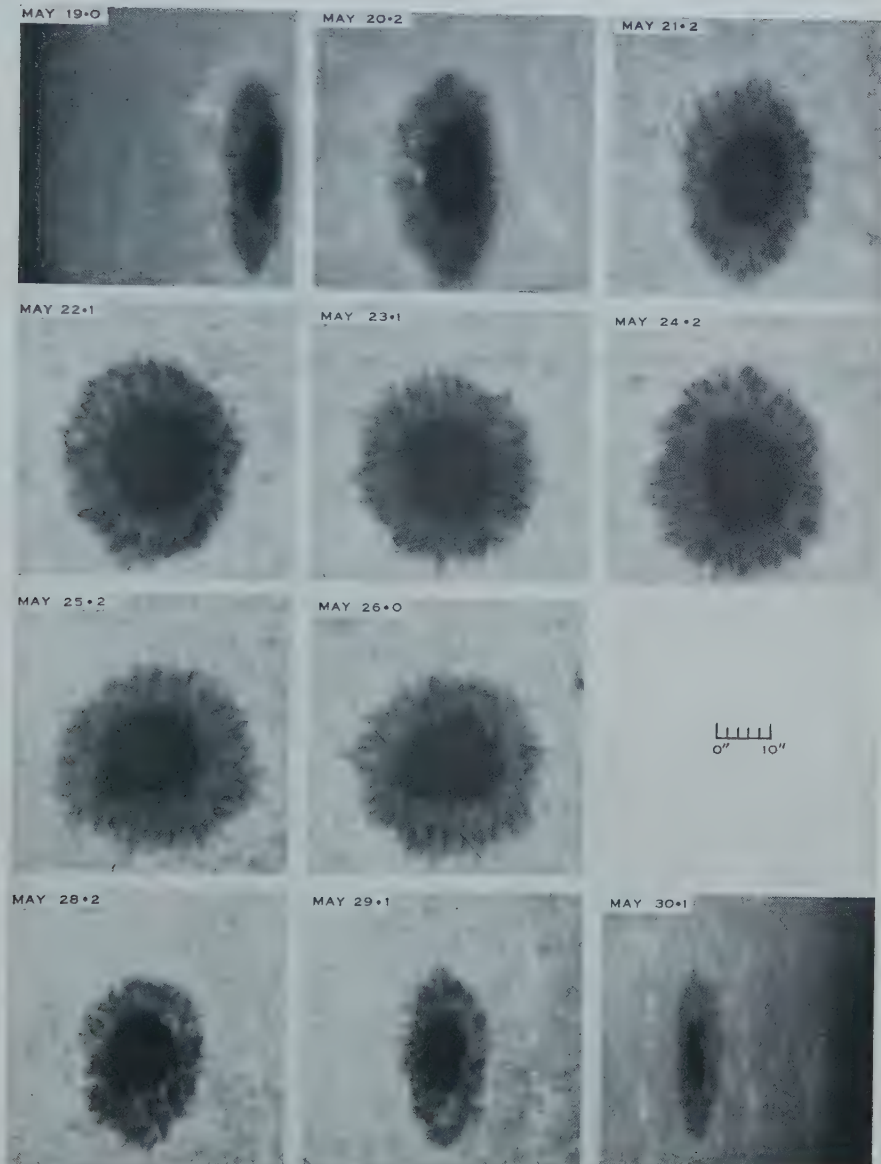
	Centre		West Limb		
Date θ	24.v.57 8.6°	25.v.57 12.1°	28.v.57 52.2°	29.v.57 64.2°	30.v.57 76.7°
AN	$7.9 \pm 0.2''$	$9.0 \pm 0.2''$	$6.1 \pm 0.2''$	$4.8 \pm 0.1''$	$2.7 \pm 0.1''$
$A'N'$	$9.7 \pm 0.4''$	$7.4 \pm 0.2''$	$5.5 \pm 0.2''$	$3.5 \pm 0.1''$	$1.8 \pm 0.1''$
f	0.82 ± 0.04	1.22 ± 0.04	1.10 ± 0.06	1.38 ± 0.07	1.51 ± 0.11

* θ is the heliocentric angle of the spot; AN , $A'N'$ are the apparent penumbral widths on the sides of the spot directed towards and away from the solar limb respectively, and $f = AN/A'N'$. It should be noted that the AN of Table 1 (a) corresponds to the $A'N'$ of Table 1 (b), and vice versa. The r.m.s. errors of the mean values AN and $A'N'$ are calculated from the formula $s = \sqrt{[\sum (\text{residuals})^2 / \{n(n-1)\}]}$, where n is the number of determinations; the corresponding errors in f are derived from these errors in the usual way. The east and west limb measurements were made along different lines in the spot, the angle between the lines being approximately 7° .

then transferred to the centre photographs and the measurements repeated. Errors due to seeing have been reduced by taking the mean of the five measures for each day (three for May 19). As experience has shown, even on a photograph of good quality, seeing can appreciably distort structures many times larger than the smallest detail resolvable on the photograph.

Tables 1(a) and 1(b) give the results of the measurements (expressed in seconds of arc) for the east and west limbs respectively, as well as the corresponding

THE WILSON EFFECTS IN SUNSPOTS



The changing appearance of the sunspot during its passage from the east to the west limb. Each photograph is orientated so that the direction of greatest foreshortening is horizontal. It will be noticed that, when the spot is near the limb, the apparent width of the penumbra on the side remote from the limb is less than that on the limb side. The heliographic coordinates on May 24 were 10°S. , 204°W.

values of the ratio f defined by equation (1). The r.m.s. errors of the mean values are also given. However, these only represent the internal consistency of the measurements and systematic error may also be present. It will be noticed that the values of f steadily increase towards either limb by amounts much greater than the r.m.s. errors.

IV. DISCUSSION

The reality of the Wilson effect is demonstrated by Plate 1.

The photographs for May 19, 20, and 21 show the appearance of the spot near the east limb. On May 19 it is seen that the width of the penumbra on the limb side is nearly double that on the other side, while the same disparity is present, though to a lesser degree, on May 20.* By May 21 the ratio of the penumbral widths is approaching unity and thereafter, until May 26, the spot displays a fairly regular structure as it moves across the central portion of the solar disk. The photographs for May 28, 29, and 30 show the appearance of the spot near the west limb. On May 29 the width of the penumbra on the limb side is again greater than that on the other side and this disparity is even more marked on May 30. However, as shown by a comparison of the results in Tables 1(a) and 1(b), the effect here is less marked than at the east limb. This reduction in the value of f may reflect the fact that the penumbral widths in the two cases are measured along somewhat different lines in the spot which, however, differ only by about 7° (cf. Section III). On the other hand it may be connected with the decay of the spot which occurs during its passage across the disk.

The fact that the Wilson effect is observed at both limbs precludes the possibility that the changes in the relative penumbral widths at either limb are due to a persistent elongation of the umbra on one side of the spot. For, since the two fixed lines of measurement in the spot differ by only 7° , the effect of any such asymmetry would be reversed at the two limbs and the results would be correspondingly discordant. The substantial agreement between the results obtained at the two limbs confirms the reality of the effect.

It is also possible that the observed effect could be due to short-lived elongations of the umbra in the direction of measurement, which develop and decay during a period of a few days or less. However, in order to explain the appearance of the limb photographs any such distortion would have to extend over at least a quarter of the umbral boundary. No such distortion can be seen on the photographs for May 21 or 22, although on the 23rd the umbra-penumbra boundary shows some irregularities. The east limb results can therefore be accepted with confidence. Such a distortion does in fact develop in the umbra on May 24. It so happens that the lines fixed in the spot corresponding to the directions of measurement at the two limbs pass through this region. However, although the distortion had increased even further by May 25, it appears to be declining on May 26. The change in this region is reflected in the discordance of the measurements for May 24 and 25 (cf. Tables 1(a) and 1(b)). But it seems

* The two small bright regions bordering the left-hand side of the umbra on this photograph are phenomena of the type recently described elsewhere (Bray and Loughhead 1957).

unlikely that the projection, even if it had continued until the spot reached the west limb on May 30, could account for the magnitude of the observed effect. In fact the measurements show that the reduction in the penumbral widths on the side of the spot remote from the limb due to this projection is probably insufficient to account for the observed effect.

V. REMARKS ON INTERPRETATION

Although there has been controversy about the reality of the effect, previous observations have generally been explained on the hypothesis that sunspots are shallow, saucer-shaped depressions in the photosphere (Young 1895; Menzel 1949; Cowling 1953, cf. pp. 569-72, 576, 578; Kiepenheuer 1953; Abetti 1957, cf. pp. 65-7, 245). However, such a geometrical interpretation is inadequate, since the whole question of the changing appearance of a sunspot during its passage across the disk is essentially a problem in the theory of radiative transfer. The principal factor in causing the Wilson effect is the greater transparency of the spot compared with the surrounding photosphere (Michard 1953; Sweet 1955), so that a line-of-sight from the observer penetrates to a greater depth in the spot than in the surrounding photosphere. Accordingly, the spot appears as a depression, whose depth depends on the heliocentric angle and on the relative transparencies of the spot and the photosphere.

Theoretically, a knowledge of the emission and absorption coefficients in the photosphere, umbra, and penumbra would yield the intensity profile of the spot along the line AA' (Fig. 1), and hence the Wilson effect. Conversely, measurements of the intensity profile at different heliocentric angles might provide some information about the emission and absorption coefficients in the spot and their variation with depth. However, any attempt at a detailed comparison of theory and observation would face serious difficulties, both observational and theoretical. The observations would require correction for scattered light and for distortion by the combined instrumental profile of the telescope and atmosphere. These corrections become very large near the limb, where the width of the foreshortened penumbra may be less than 2 sec of arc. On the theoretical side there is the difficulty of matching the scales of optical depth in the spot and photosphere respectively (Michard 1953, cf. p. 280; Sweet 1955).

Some insight into the cause of the Wilson effect can be obtained by considering a simple model in which the sunspot is represented as a cylindrical structure extending through the photosphere, and in which the absorption coefficients κ_1 , κ_2 , κ_3 of the photosphere, penumbra, and umbra respectively are constant (Fig. 2). The photosphere and the spot are assumed to have a sharp upper boundary. Making the reasonable assumption that $\kappa_1 > \kappa_2 > \kappa_3$ and taking the directions of the apparent photosphere-penumbra and umbra-penumbra boundaries to be such that the corresponding lines-of-sight pass through unit optical thickness, it is easy to show that the apparent width of the penumbra on the side of the spot remote from the limb decreases more rapidly than a geometrical foreshortening law would imply, while on the other side of the spot the reverse applies. This is the Wilson effect. Moreover, in agreement with observation (Waldmeier 1955, cf. p. 163), the area of the whole spot decreases faster than

implied by geometrical foreshortening. This simple model therefore explains qualitatively not only the Wilson effect but also the anomalous foreshortening of sunspot areas.* However, values of κ_2 and κ_3 obtained from the observations on the basis of this model probably have little meaning. The observations indicate that in this particular sunspot the Wilson effect is due mainly to the

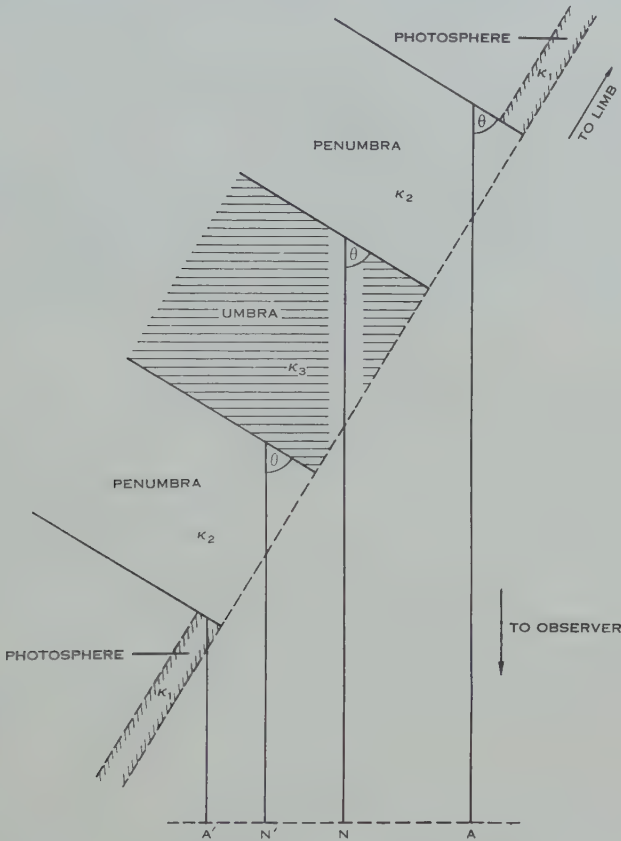


Fig. 2.—A simple model of a sunspot. The figure shows a section of the spot by the plane of the great circle containing the solar radius and the line-of-sight to the spot.

reduced foreshortening on the limb side of the spot and not to the increased foreshortening on the other side. If this is true for spots in general then the model implies that the dominant factor is the great transparency of the umbra, rather than that of the penumbra, compared with that of the photosphere.

VI. ACKNOWLEDGMENTS

The authors wish to thank Mr. V. R. Burgess for assistance in securing some of the observations and Mr. H. Gillett for processing the films and making the enlargements.

* Hotinli (1957) attributes both the Wilson effect and the failure of the cosine law to an elevated ring of faculae surrounding the spot.

VII. REFERENCES

- ABBOT, C. G. (1912).—"The Sun." p. 199. (Appleton: London and New York.)
- ABETTI, G. (1957).—"The Sun." (Faber and Faber: London.)
- BRAY, R. J., and LOUGHHEAD, R. E. (1957).—*Observatory* **77**: 201.
- COWLING, T. G. (1953).—"The Sun." (Ed. G. Kuiper.) Ch. 8. (Univ. Chicago Press.)
- HOTINLI, M. (1957).—*Ann. Astrophys.* **20**: 45.
- KIEPENHEUER, K. O. (1953).—"The Sun." (Ed. G. Kuiper.) pp. 324, 330. (Univ. Chicago Press.)
- LOUGHHEAD, R. E., and BURGESS, V. R. (1958).—*Aust. J. Phys.* **11**: 35.
- MENZEL, D. H. (1949).—"Our Sun." pp. 124-5. (Churchill: London.)
- MICHARD, R. (1953).—*Ann. Astrophys.* **16**: 217.
- SWEET, P. A. (1955).—"Vistas in Astronomy." Vol. I. (Ed. A. Beer.) p. 677. (Pergamon: London and New York.)
- UNSÖLD, A. (1955).—"Physik der Sternatmosphären." 2nd Ed. p. 559. (Springer: Berlin.)
- WALDMEIER, M. (1955).—"Ergebnisse und Probleme der Sonnenforschung." 2nd Ed. (Geest u. Portig: Leipzig.)
- YOUNG, C. A. (1895).—"The Sun." pp. 128-31. (Kegan Paul: London.)

THE LIFETIME OF SUNSPOT PENUMBRA FILAMENTS

By R. J. BRAY* and R. E. LOUGHHEAD*

[*Manuscript received December 23, 1957*]

Summary

Twenty-seven photographs of a sunspot showing a particularly distinct filamentary structure are analysed in order to obtain an estimate of the lifetime of the penumbral filaments. The photographs have been selected from some 2400 taken over a period of 5 hr. Although the measurements are rendered somewhat uncertain by observational difficulties the results clearly indicate that some filaments remain identifiable for periods of the order of hours.

I. INTRODUCTION

No measurements of the lifetime of the filaments observed in the penumbrae of sunspots have hitherto been published, though various authors (Kiepenheuer 1953 ; Macris 1953 ; Waldmeier 1955) have stated that it exceeds the lifetime of the photospheric granulation, which is of the order of a few minutes. In this paper an account is given of a systematic attempt to measure the lifetime of penumbral filaments in a sunspot photographed near the time of its maximum development. Although the measurements are rendered somewhat uncertain by the narrow width, low contrast, and confused structure of the filaments, the results clearly indicate that some filaments remain identifiable for periods of the order of hours. It is not known, however, whether these results are typical of other spots at other stages of development.

II. OBSERVATIONS

The observations were made with the 5 in photoheliograph (Loughhead and Burgess 1958) of the C.S.I.R.O. Division of Physics Solar Observatory. This instrument is designed to photograph any selected region of the solar disk on 35 mm film at 5-sec intervals; the diameter of the solar image is 20 cm. Plate 1 (*a*) shows the spot selected for study; this particular spot was distinguished from others contained in the film records of the Observatory by possessing an unusually clear filamentary structure. Even so, the structure is not equally distinct all round the penumbra. Measurements were therefore restricted to the lower, right-hand quadrant, where the filaments are clearest. An enlargement of this region, made with enhanced contrast, is shown in Plate 1 (*b*).

Even on photographs of the quality of Plate 1, where the filamentary structure of the penumbra is well resolved, it is extremely difficult to identify individual filaments on successive photographs. The reason for this is that the structure is complex, the contrast between the bright filaments and the darker

* Division of Physics, C.S.I.R.O., University Grounds, Chippendale, N.S.W.

penumbral material is low, and the filaments are very narrow. In most cases the apparent width of the filament lies between 0.5 and 1 sec of arc, so that the detail is even finer than the photospheric granulation and is, in fact, comparable with, or less than, the theoretical resolving limit of the telescope (0.8 sec of arc). Consequently, the structure of the filaments can be seriously distorted by atmospheric seeing even on photographs on which the granulation is well resolved. For example, a filament which is seen on a given photograph as a continuous line might appear on a subsequent photograph, taken only a few seconds later, broken into two or more segments. Similarly, two neighbouring filaments might appear similar in shape on adjacent photographs, but with a different separation; or one filament might be completely obliterated, although its neighbour remains intact. For these reasons it is impossible to follow in detail any structural changes undergone by individual filaments. However, as shown below, it is possible to estimate their lifetimes.

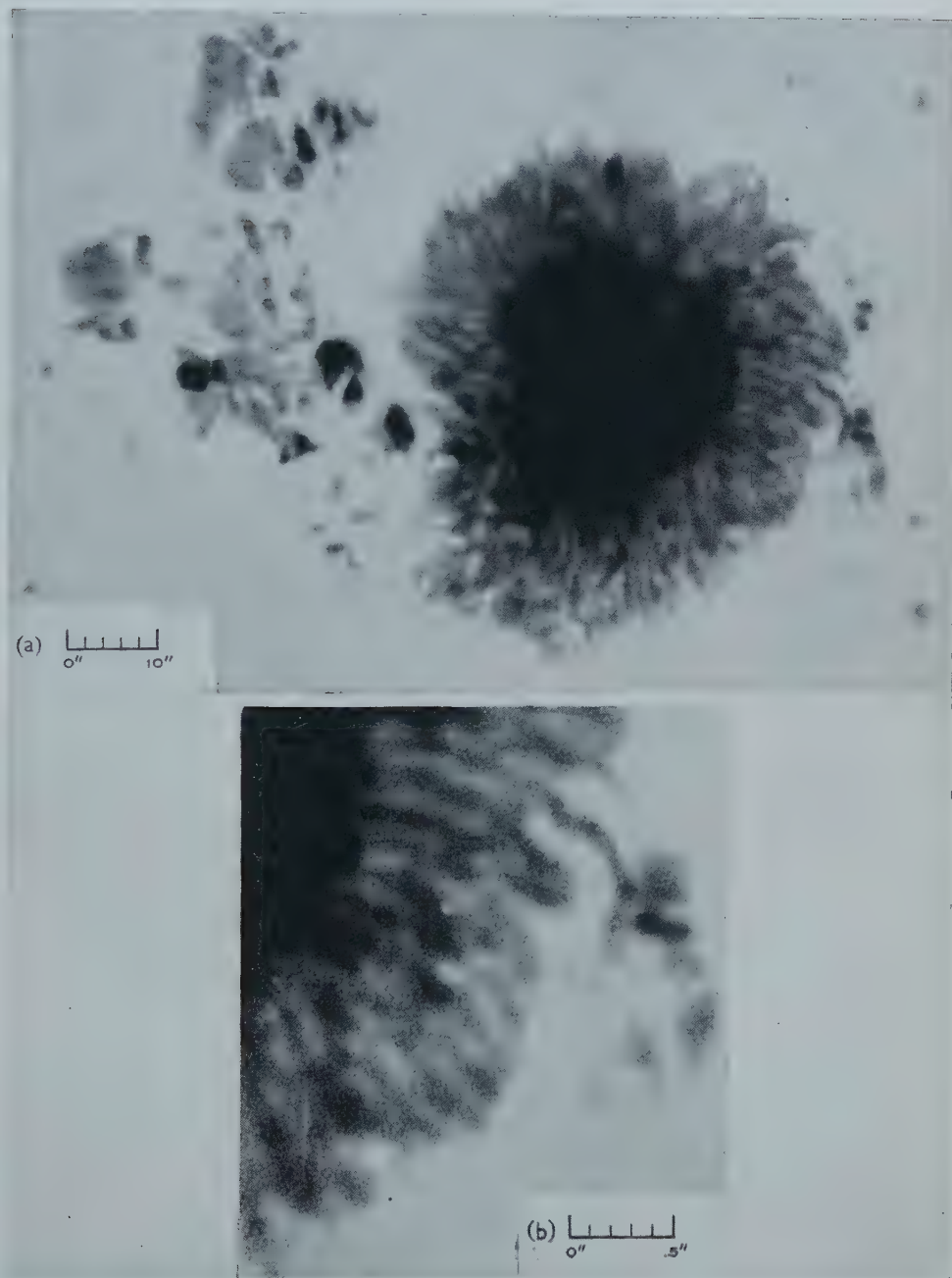
The spot shown in Plate 1 was photographed over a period of 5 hr on June 7, 1957. Although none of the 2400 odd photographs on the film equals the best hitherto obtained with the instrument, the film contains 27 good photographs distributed fairly uniformly over a period of about 200 min. The measurements described in the next section were made from enlargements of these photographs, which are on such a scale that 1 mm corresponds to 0.9 sec of arc.

III. REDUCTION

In view of the difficulty of identifying individual filaments on successive photographs it was decided to make maps of the filaments for each of the 27 photographs. This procedure introduces a considerable simplification without losing the essential features of the structure. The maps were restricted to the region of the spot shown in Plate 1 (*b*), where the filamentary structure is particularly clear. In tracing out the filaments a gap was left where there intervened either a dark region or a bright diffuse patch of the type described by Macris (1953). All maps were drawn independently, without reference to one another. Also, to facilitate the intercomparison of the maps, the outlines of the umbra and of a few small adjacent spots were included. However, owing to displacements (~ 1 sec of arc) due to seeing, the final matching of any two maps could only be made by slightly displacing one until the best fit was obtained for the filamentary detail.

To systematize the intercomparison of the 27 maps a "key" map was first selected: from the 27 maps, 13 were chosen which appeared to show the clearest filamentary structure; these evidently represent the 13 photographs least affected by seeing. Each of the 13 maps was then compared in turn with the remaining 12, and the degree of correlation in each case subjectively assessed on a scale of 1 (good) to 3 (poor). The number of correlations of class 1 for each map was adopted as a figure of merit for this map. A low value for this figure indicates that the corresponding photograph is affected by seeing; a high value indicates that it closely represents the structure under study. It is interesting to note that, although on visual inspection the 13 photographs appear to be of comparable

THE LIFETIME OF SUNSPOT PENUMBRA FILAMENTS



(a) Appearance of spot at 11^h 58^m, June 7, 1957. Heliographic coordinates: 17°S., 8°W.

(b) The region of the penumbra selected for study; the enlargement was made with enhanced contrast.

quality, figures of merit ranging from 1 to 9 were obtained. The map having the highest value for this figure was selected as the key and is shown in Figure 1.* Six particularly distinct filaments on this map are numbered.

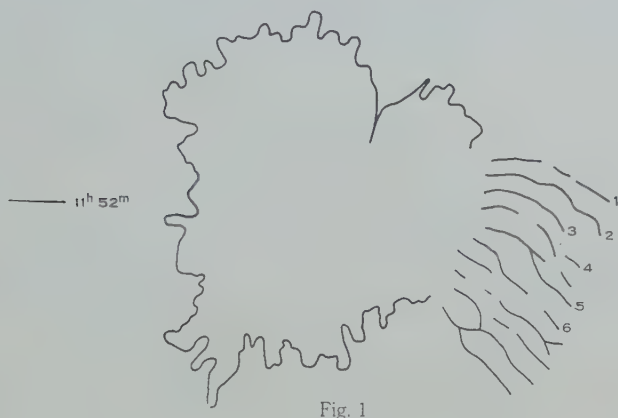


Fig. 1

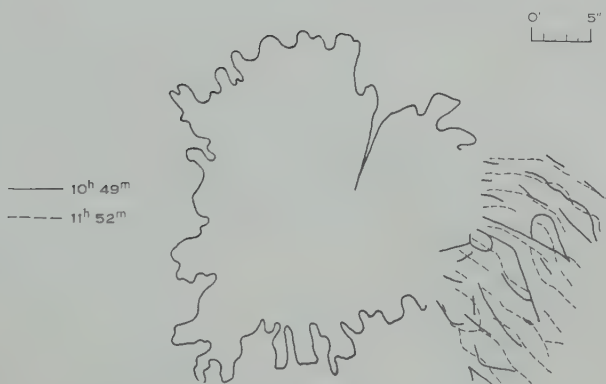


Fig. 2

Fig. 1—Penumbral structure at $11^{\text{h}} 52^{\text{m}}$. This figure is the “key” map (see text). In Figures 2–6, the structure at $11^{\text{h}} 52^{\text{m}}$ (shown dotted) is compared with the structure at earlier and later times respectively. Note that the umbral outlines correspond to the times of the figures.

In Figures 2–6 the key map is compared with five other maps corresponding to earlier and later times. At $10^{\text{h}} 49^{\text{m}}$ (Fig. 2), 63 min before the time of the key map, the correlation is weak, only filament 2 being identifiable. At $11^{\text{h}} 8^{\text{m}}$ (Fig. 3), 44 min before the key, the correlation is somewhat stronger, filaments

* This method of selecting the most representative of a number of photographs, all apparently affected by atmospheric seeing more or less equally, is essentially an *autocorrelation technique*. It is clearly only applicable if the interval between successive photographs is significantly shorter than the average lifetime of the features under study. The method favours good photographs occurring near the middle of the sequence. It could be useful in determining the lifetime of other solar features, e.g. the photospheric granulation.

1, 2, 3, 5, and 6 being identifiable. At 11^h 58^m (Fig. 4), only 6 min after the key, the correlation, as one would expect, is very marked; the close agreement shown here confirms the reliability of the procedure adopted. At 12^h 39^m (Fig. 5), 47 min after the key, filaments 1, 2, and 3 are identifiable. At 13^h 12^m (Fig. 6), 80 min after the key, the correlation is much reduced. Figures 2-6 indicate that certain filaments remain identifiable for periods of the order of hours.

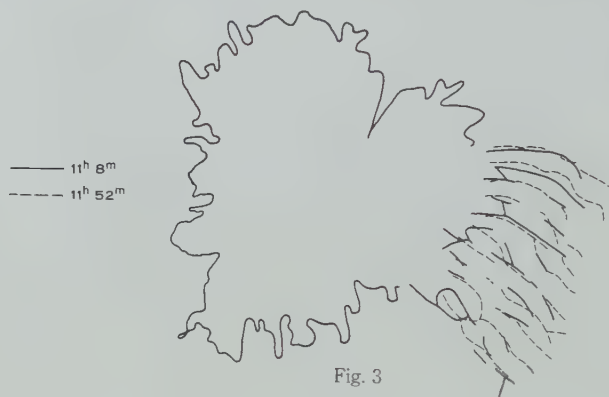


Fig. 3



Fig. 4

Estimates of the lifetimes of the 6 particularly distinct filaments labelled in Figure 1 were obtained by a comparison of the key map with each of the 26 remaining maps. In carrying out this comparison, a given filament was arbitrarily taken to be present if it had the same shape, occurred at nearly the same position, and occupied at least 50 per cent. of the length of the same filament shown on the key map. With this criterion the following results were obtained: filament No. 1 can be identified on 12 maps extending from 11^h 7^m to 12^h 39^m; two maps during this period fail to show it. No. 2 is visible on 18 maps extending from 10^h 24^m to 14^h 21^m; 7 maps during this period fail to show it. No. 3 is visible on 10 maps extending from 11^h 7^m to 12^h 39^m; 4 maps during this period fail to show it.

Filaments Nos. 4, 5, and 6 are rather uncertain. They are visible on a photograph taken near $11^{\text{h}} 8^{\text{m}}$, but several other photographs taken within a minute fail to show them. No. 4 is only certainly present from $11^{\text{h}} 52^{\text{m}}$ to $11^{\text{h}} 58^{\text{m}}$ (3 maps). No. 5 is present from $11^{\text{h}} 52^{\text{m}}$ to $12^{\text{h}} 39^{\text{m}}$ (5 maps); 3 maps during this period fail to show it. No. 6 is visible on 3 maps extending from $11^{\text{h}} 52^{\text{m}}$ to $12^{\text{h}} 16^{\text{m}}$; 1 map during this period fails to show it.

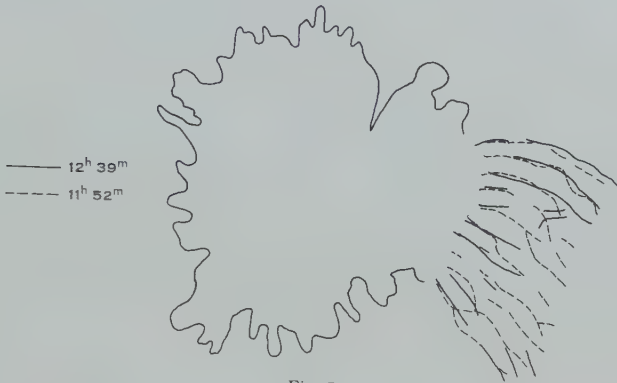


Fig. 5



Fig. 6

The results of the measurements may be summarized by stating that values ranging up to several hours have been found for the lifetimes of the six filaments which were distinct enough for a determination to be made.

IV. DISCUSSION

Observations of the lifetime of penumbral filaments may be expected to throw some light on the question of the dynamical stability of sunspots and, in particular, on the Evershed effect. The simplest interpretation of the Evershed effect is that it consists of a laminar flow of matter outwards from the umbra along the filaments, which are probably shallow structures of depth comparable with their width. With an average filament length of 7500 km for the spot under

study and a mean Evershed velocity of 1 km/sec (Kinman 1953) the time taken by matter to flow along the entire length of a filament is about 2 hr. It may be significant that this figure is of the same order of magnitude as some of the observed lifetimes.

However, the observations of St. John (1913*a*, 1913*b*), who found that the Evershed velocity decreases with height, ultimately reversing its direction, throw doubt on this simple interpretation of the Evershed effect. St. John deduced his result from measurement of strong Fraunhofer lines, whose cores are formed higher in the Sun's atmosphere than those of the weaker lines normally used. However, the effective level of observation in the case of a strong Fraunhofer line observed with a spectrograph of moderate resolving power may differ from that of a weak line by only a few tens of kilometres (cf. Hart 1956). This is due to the effect of the outer parts of the instrumental profile, which throw light from the wings of a Fraunhofer line into the core. St. John's results therefore indicate that the Evershed velocity may change appreciably over a very small depth. In view of the great length of the filaments compared with their width it is hard to understand how they can maintain such a steep velocity gradient without disintegration.

V. ACKNOWLEDGMENT

The authors wish to thank Mr. H. Gillett for processing the films and making the enlargements.

VI. REFERENCES

- HART, A. B. (1956).—*Mon. Not. R. Astr. Soc.* **116**: 489.
KIEPENHEUER, K. O. (1953).—"The Sun." (Ed. G. Kuiper.) p. 346. (Univ. Chicago Press.)
KINMAN, T. D. (1953).—*Mon. Not. R. Astr. Soc.* **113**: 613.
LOUGHHEAD, R. E., and BURGESS, V. R. (1958).—*Aust. J. Phys.* **11**: 35.
MACRIS, C. (1953).—*Ann. Astrophys.* **16**: 19.
ST. JOHN, C. E. (1913*a*).—*Astrophys. J.* **37**: 322.
ST. JOHN, C. E. (1913*b*).—*Astrophys. J.* **38**: 341.
WALDMEIER, M. (1955).—"Ergebnisse und Probleme der Sonnenforschung." 2nd Ed. p. 185. (Geest u. Portig: Leipzig.)

THE FLARE-SURGE EVENT

By R. G. GIOVANELLI* and MARIE K. MCCABE*

[*Manuscript received January 21, 1958*]

Summary

Studies of flare-surges recorded on low magnification $H\alpha$ flare patrol films show that dark surges on the disk come directly from flares, the initial direction of motion being almost invariably away from the nearest large sunspot.

The basic feature of the event is the emission of a substantially continuous particle stream having different appearances at different stages. The first trace is usually the ejection from the flare of diffuse matter, brighter than the chromosphere, which fades and disappears by becoming transparent, though it can be detected if the stream crosses the limb. The dark surge subsequently appears at the base of the stream, close to the flare.

Dark surges, bright streamers on the disk, and ejected flares seem to be closely related phenomena. It is shown that dark surges can develop only if the temperature is decreasing, while in ejected flares whose brightness increases during flight the temperature must be increasing.

I. INTRODUCTION

The term "surge" is usually applied to ejections of matter from the solar chromospheric regions, observable on the disk or at the limb because of significant absorption or emission in $H\alpha$ and certain other spectral lines; the term does not include ejections arising from pre-existing filaments.

The occurrence of surges at times of flares has often been noted, but one vital aspect of the association has remained obscure, namely, the location of the surge with respect to the flare. The main papers on surges agree that, when first observed, the surge appears near but not usually in contact with the flare. The question arises whether the surge comes from the flare or whether it is an allied event occurring simultaneously in another part of the active region and having a cause in common with the flare. We propose to describe observations which settle this matter and enable a clearer description of the flare-surge event to be given.

The surges which accompany some flares, appearing dark against the chromospheric background when viewed with a suitable $H\alpha$ monochromatic telescope, had been studied first by Giovanelli (1940), Ellison (1942), Newton (1942), and Bruzek (1951). From visual observations they had found that surges were associated with about one-third of the flares of class 1 or greater, being ejected with velocities of 100 km/sec or more; and that when first observed these surges were usually displaced from the flare by a gap of some 10^4 – 10^5 km.

* Division of Physics, C.S.I.R.O., University Grounds, Chippendale, N.S.W.

It is well known that flares themselves do not generally show velocities of such magnitudes, though Giovanelli (*loc. cit.*) had already observed in 1939 a limb flare rising at 78 km/sec. Severny and Shaposhnikova (1954) have pointed out, however, that in some flares there are motions of matter in the shape of streaks and ejections. Earlier, Dodson and Hedeman (1949) had recorded the ejection from several flares of long streamers, initially brighter than the undisturbed chromosphere, but subsequently becoming dark; they suggested that some flares are the root points or bases of surges, and in particular that the bright streamer on the disk is the counterpart of the rarer limb surges which are more brilliant than usual. Dodson and McMath (1952) drew attention to some exceptionally bright surge-type prominences which, they said, might or might not be the limb aspects of features that would appear as small flares or flare-like brightenings if seen on the disk. Also they reported a very small number of major prominence phenomena believed to be the limb aspects of important flares; in one case, portion of a very brilliant formation rose with a velocity of 700 km/sec to a height of 50,000 km. Dodson and Hedeman (1952) described the presence of a short-lived prominence, bright in projection against the disk, after a flare. More recently, Bray *et al.* (1957) noted the ejection of a bright flare on the disk, at a velocity of 300 km/sec. This severed contact with a stationary base flare and reached maximum brightness at a height of 25,000 km above the chromosphere. This ejection later crossed the limb, where its appearance resembled an ordinary surge.

The observations described below reveal no new phenomena, but enable us to obtain a coherent picture of the flare-surge event. In particular, it appears that dark surges observed on the disk come directly from the flares; but almost invariably, at least when the resolution permits it, the first trace of the surge is the ejection of diffuse matter, brighter than the chromosphere. The stream of ejected matter generally fades and disappears by becoming transparent, though the ejection of invisible material continues. Subsequently, absorbing matter appears in the stream, either superimposed on or close to the flare, and travels outwards to appear as the normal dark surge. The ejection of high speed flares of the type observed by Bray *et al.* is interpreted as an extreme case of the flare-surge event.

II. OBSERVATIONS

The Sydney Lyot-monochromator, with a bandwidth of 0.7 \AA centred on $H\alpha$, is used in conjunction with a 5 in. telescope to obtain, for "flare patrol" purposes, 16 mm diameter photographs of the solar chromosphere on 35 mm Eastman Kodak IV E film at 0.5 min intervals. While these records lack much of the resolution obtainable with higher initial magnification, they have enabled us to observe a great number of flares over the past 2 years. We believe that most flares occurring during the time of observation have been recorded if their areas exceeded about 2×10^{-5} of the Sun's hemisphere. Surges are also observable, though not necessarily with maximum contrast, since line-of-sight velocities shift the centre of the absorption line away from the wavelength of maximum transmission.

The present discussion is based on observations from November 9, 1955 to July 4, 1957. The numerical analyses are based on observations from November 9, 1955 to August 16, 1956, and November 5, 1956 to December 12, 1956, though not all flares observed during these intervals have been included. For convenience, before July 23, 1956 our study has been confined to periods when simultaneous radio-frequency spectrum observations were being made at Dapto, N.S.W. For the subsequent period flares have been included only if they occurred within about 45° of the central meridian. These two sets cover 482 flares, the majority being very small, of class 1—. The distribution of apparent flare areas is given in the histogram of Figure 1; the drop for areas

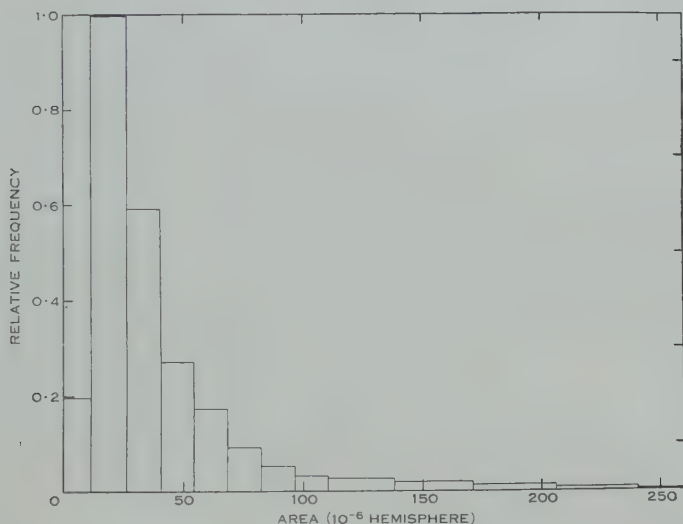


Fig. 1.—Distribution of apparent areas of 482 flares (millionths of the Sun's hemisphere) observed in Sydney during the period described in Section II.

below 1.4×10^{-5} of the Sun's hemisphere is probably due to the limits of resolution. Of the above flares, 78 were accompanied by dark surges on the disk, some large, some very small. In 46 other cases, the observations were inadequate to decide; deducting these from the total, we find that 18 per cent. of the flares were associated with surges. There was no statistically significant variation with flare size.

During this whole period we have discovered no dark surges which were not associated with flares. We conclude that all dark surges are associated with flares, most of which are of class 1—.

Plates 1-4 and Plate 5, Figure 1, show a number of flares and associated surges which are typical of the general range of surge appearance. Descriptions of interesting features are given below; times are U.T.

January 24, 1957

This flare developed in a region with only a very small sunspot group. The flare expanded rapidly northwards until $00^{\text{h}} 53^{\text{m}} 15^{\text{s}}$. A minute later, traces of

a dark surge appeared superimposed on the flare; as the flare faded, the surge gradually extended outwards in the same direction as the bright expansion.

July 4, 1957

A bigger flare, this time near a larger sunspot, showing a similar sequence of events. An initial diffuse bright expansion of the flare occurred away from the sunspot, with maximum development at $00^{\text{h}} 33^{\text{m}} 0^{\text{s}}$. The diffuse part faded rapidly. A dark surge appeared in the same region at $00^{\text{h}} 37^{\text{m}} 0^{\text{s}}$, and expanded radially away from the sunspot.

November 26, 1956

One of a number of fine flare-surges associated with this sunspot. The flare, expanding away from the sunspot, had a very diffuse outer edge. The major dark surge can be traced back to the appearance of absorbing matter superimposed on the flare at $01^{\text{h}} 08^{\text{m}} 30^{\text{s}}$; this gradually expanded outwards until about $01^{\text{h}} 50^{\text{m}}$. In its earlier stages the dark surge appeared fairly transparent—it is possible to detect underlying structure through it—though it would seem to have been fairly opaque over most of its length at $01^{\text{h}} 29^{\text{m}} 30^{\text{s}}$. Faint bright extensions of the flare were present on either side of the dark surge, e.g. at $01^{\text{h}} 22^{\text{m}}$, but these gave place to the much wider dark surge later.

The stream of ejected matter seemed to have been in existence during most of the event. It was never of very high opacity; the substantial changes in appearance during the event (e.g. from $01^{\text{h}} 25^{\text{m}} 30^{\text{s}}$) seem to have been due more to changes in excitation and opacity than to major changes in overall dimensions.

May 18, 1956

The flare, beginning at $01^{\text{h}} 23^{\text{m}} 52^{\text{s}}$, showed a rapid bright expansion away from its neighbouring sunspot, starting at $01^{\text{h}} 29^{\text{m}} 22^{\text{s}}$ and with maximum extent at $01^{\text{h}} 30^{\text{m}} 52^{\text{s}}$. It then faded back from the tip of the flare until the appearance at $01^{\text{h}} 36^{\text{m}} 52^{\text{s}}$ (not shown) resembled that shortly after commencement. At $01^{\text{h}} 37^{\text{m}} 52^{\text{s}}$ (not shown), however, a small spike appeared on enhanced exposures made at the limb. A dark surge also made its first appearance at this time in the position formerly occupied by the bright flare expansion. This dark surge moved out, joining the limb spike at about $01^{\text{h}} 47^{\text{m}} 52^{\text{s}}$; the continuous outward development of this spike demonstrated the continuity of the stream of ejected matter throughout the event.

The time between the initial flare expansion at $01^{\text{h}} 29^{\text{m}} 22^{\text{s}}$ and the first appearance of the limb spike at $01^{\text{h}} 37^{\text{m}} 52^{\text{s}}$ was about the same as that taken by the dark surge to reach the limb, so that the average velocity of the stream was approximately constant over this period, while a time plot of the position of the tip of the dark surge shows that the transverse velocity was effectively constant along the stream. Now, except perhaps just after its first appearance, the dark surge was separated from the flare, with the chromosphere apparently visible between; certainly the chromosphere was plainly visible in the later stages. The continuous increase in length of the dark surge therefore seems

to have been due to the continual formation of new absorbing material at the rear of the dark surge; in other words, an invisible stream of matter ejected from the flare became opaque.

The resolution was not sufficient to establish whether the same material forming the base of the diffuse bright stream at $01^{\text{h}} 34^{\text{m}} 52^{\text{s}}$ formed the tip of the dark surge further out at $01^{\text{h}} 38^{\text{m}} 52^{\text{s}}$.

August 24, 1956

This was rather similar to the event of May 18, 1956. A flare expanded away from a neighbouring sunspot, having maximum extension at about $04^{\text{h}} 09^{\text{m}} 35^{\text{s}}$. A minute later, an enhanced exposure revealed a small projection beyond the limb, though the high transparency of the stream was indicated by its non-appearance immediately within the limb. The bright stream faded, particularly near the flare, leaving a distant brightish part near the limb at $04^{\text{h}} 13^{\text{m}} 35^{\text{s}}$. In the meantime a dark surge developed on the upper side of the stream. At about $04^{\text{h}} 12^{\text{m}} 35^{\text{s}}$ it extended out to the limb. Even allowing for the movement along the stream, it seems very likely that some of the originally bright matter at the base of the stream at $04^{\text{h}} 10^{\text{m}} 35^{\text{s}}$ had moved out and become dark by $04^{\text{h}} 12^{\text{m}} 35^{\text{s}}$. The enhanced limb photograph at $04^{\text{h}} 23^{\text{m}} 35^{\text{s}}$ showed that the particle stream was much more extensive than indicated by its appearance on the disk, providing clear evidence for the high transparency of most of the stream.

January 18, 1957

Immediately after its outbreak this flare showed a bright diffuse expansion away from the small neighbouring sunspot, reaching maximum extent at $04^{\text{h}} 03^{\text{m}} 12^{\text{s}}$. A dark surge appeared first at $04^{\text{h}} 06^{\text{m}} 12^{\text{s}}$, in contact with the flare. From $04^{\text{h}} 07^{\text{m}} 12^{\text{s}}$ the rate of expansion of the surge was approximately constant, much less than the apparent rate between $04^{\text{h}} 06^{\text{m}} 12^{\text{s}}$ and $04^{\text{h}} 07^{\text{m}} 12^{\text{s}}$; furthermore, backward extrapolation of the rate of growth indicates that the material was ejected initially at the time of bright expansion. This is a clear case where the dark surge was formed by a pre-existing invisible particle stream becoming opaque. It is uncertain whether any of the material participating in the bright diffuse expansion of the flare at $04^{\text{h}} 03^{\text{m}} 12^{\text{s}}$ was identical with that forming the top of the dark surge, but it seems very likely.

May 30, 1956

In this event a bright streamer was ejected from a tiny flare near the limb. The bright streamer crossed the limb, beyond which its faintness showed that it was fairly transparent. An enhanced limb exposure at $00^{\text{h}} 51^{\text{m}} 35^{\text{s}}$ showed a surge, providing evidence of the diffuseness of the edges of the ejected stream.

May 7, 1957

This is a rare type of ejection, similar to that described by Bray *et al.* (1957), in which part of the flare was itself ejected. By $02^{\text{h}} 38^{\text{m}} 37^{\text{s}}$ no trace of the ejection could be seen on the disk; it had become quite transparent. However, an enhanced exposure half a minute later revealed a surge beyond the limb,

from whose size it is clear that the stream was more extensive than appeared on the disk. At $02^{\text{h}} 39^{\text{m}} 07^{\text{s}}$ first traces of a dark surge were seen in contact with the stationary flare. This travelled out along the same path, disappearing, however, before reaching the limb.

August 23, 1956

This event consisted of a flare in contact with a sunspot. At $03^{\text{h}} 32^{\text{m}} 15^{\text{s}}$ the flare expanded away from the sunspot and began to fade a few minutes later. About $03^{\text{h}} 36^{\text{m}} 15^{\text{s}}$ a dark marking could be seen adjacent to the western edge of the now fading flare and extending radially away from the sunspot. The subsequent frames show the development of the dark marking into a fine dark surge. Simultaneously the bright matter ejected by the flare moved further outwards, apparently in contact with but below the dark surge, and appeared, particularly at $03^{\text{h}} 39^{\text{m}} 15^{\text{s}}$, as a bright marking separate from the remnants of the flare.

III. RELATIVE POSITIONS OF FLARE AND SURGE

As illustrated by Plates 1–4 and Plate 5, Figure 1, most surges have been found to overlap or to start almost in contact with their associated flares. In cases where the surge has not originated in contact with the flare, almost invariably the direction of motion has been such that, if projected backwards, it would have crossed or touched the flare. However, as seen clearly in the events of July 4, 1957 and August 23, 1956, the surge has not always come from the centre of the flare.

In 2 cases out of 78 the surge appeared to originate near the flare but not in the flare itself.

IV. DIRECTION OF SURGE

Since our observations do not yield line of sight velocities, we are able to deal only with the projected paths of the dark surges on the Sun's disk.

In the examples already given the initial paths of surges lie generally radially away from the nearest sunspot. To see whether this is a general conclusion we have carried out two sets of analyses:

(a) November 9, 1955–June 21, 1956. Sketches have been made of the positions of 29 surges in spot groups well away from the limb, without any attempt to correct for foreshortening.

(b) June 19, 1956–December 14, 1956. The solar image has been projected onto a hemispherical screen, sketches of the positions of surges in spot groups thus being approximately corrected for foreshortening. The 48 surges involved all occurred within about 45° from the centre of the disk.

For each set the sketches have been superimposed, with the nearest large sunspot at a fixed position (Figs. 2 (a), 2 (b)). With only one or two exceptions, the initial paths of these surges are more or less radially away from the sunspot.

V. DIFFUSE BRIGHT EXPANSION PRECEDING SURGES

We have noted an interesting feature which is almost always followed by a dark surge. This is a diffuse, usually asymmetric, expansion of the flare. As the diffuse bright material fades, it gives place to the dark surge which then moves outwards. Plates 1-4 and Plate 5, Figure 1, all show this effect, though it is much more evident when the films are examined in cinematographic projection.

Of 78 flare-surge events studied on the disk, this feature could be identified in 39, with 16 doubtful cases. Discarding the latter, it appears that at least 60 per cent. of our events are of this type. Further, they include 82 per cent.

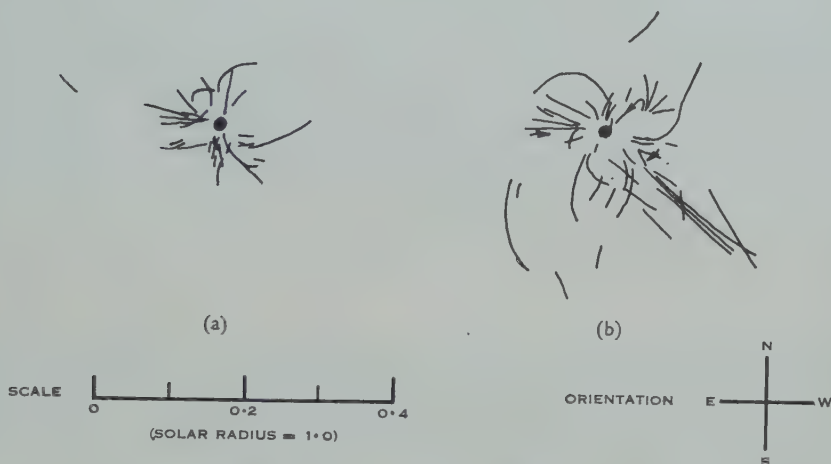


Fig. 2 (a).—Positions of 29 surges superimposed with respect to the nearest sunspot. Surges occurred between November 9, 1955 and June 21, 1956. All regions are well away from the Sun's limb, and the surges have been drawn without correction for foreshortening. Directions of surges are substantially away from the sunspot.

Fig. 2 (b).—Positions of 48 surges superimposed with respect to the nearest sunspot. Surges occurred between June 19, 1956 and December 14, 1956. The solar image has been first projected onto a hemispherical screen, sketches then being made of the positions of surge and sunspot. All surges occur within about 45° from the centre of the disk. Directions of surges are substantially away from the sunspot except where indicated by arrows.

of events involving flares of apparent area exceeding 2.8×10^{-5} of the Sun's hemisphere, as against 30 per cent. of events involving smaller flares. We believe that the difference is due solely to our inability to resolve the smaller flares adequately.

This is the type of event described first by Dodson and her collaborators. We can now state, however, that it is probably the normal mode of development of all flare-surge events.

VI. LIMB APPEARANCES OF SURGES

It is of interest now to consider the appearance of similar flare surges at the limb.

The best example we have observed occurred on the west limb on November 30, 1956, when the large spot group at 25°S. was just disappearing. This group had been remarkable in surge productivity, some of our finest examples being observed in its neighbourhood on November 26, 1956 (see Plate 2, Fig. 1). Flare surges were noted on succeeding days. On November 30, a similar type of event occurred on the limb (Plate 5, Fig. 2), where a flare broke out at $05^{\text{h}} 05^{\text{m}} 50^{\text{s}}$, rising slowly to a height of 7000 km by $05^{\text{h}} 09^{\text{m}} 50^{\text{s}}$. The edges of the flare were very diffuse, and the brightness exceeded that of the chromosphere out to a height of 28,000 km at this time. Photographs of longer exposure showed the feature to be much wider and higher. The flare faded soon after $05^{\text{h}} 09^{\text{m}} 50^{\text{s}}$, but a fine surge had developed to a height of 130,000 km by $05^{\text{h}} 19^{\text{m}} 50^{\text{s}}$.

Many spikes or surges seen at the limb would be invisible if on the disk. Not only are they of low brightness, requiring longer exposures for registration, but they are highly transparent. This is illustrated by limb surges associated with the spot group at $20^{\circ}\text{S.}, 70^{\circ}\text{W.}$ on May 30, 1956. On this day 17 distinct small flares were observed at a mean longitude 75°W. from the central meridian. Enhanced limb exposures showed that 10 of these flares emitted small surges, following the emission of faint bright matter which in most cases did not appear to extend to the limb. Only in one case did the bright matter turn dark on the disk; in all other cases the ejected surge was of very low optical thickness.

VII. THE TRANSITION FROM FLARE TO DARK SURGE

In the transition from flare to dark surge the question arises whether the ejected material passes through a completely transparent stage or through an opaque stage in which it is of the same brightness as the background. One of our difficulties in studying this point has been the small scale of the photographs, because of which there are very few events suitable for examination.

The flares of Plate 1, Figure 2; Plate 2, Figure 2; and Plates 3 and 4 show quite conclusively that the bright expansion which occurs at the beginning of the flare disappears by becoming transparent. The clearest example is that of May 18, 1956, in which enhanced limb exposures showed the ejected matter passing beyond the limb while previously invisible on the disk. This conclusion is reinforced by numerous other such examples.

The flares of November 26, 1956 and January 18, 1957 provide clear examples in which the dark surge forms as the result of an invisible cloud of ejected particles becoming opaque. In many other events, such as those of November 26, 1956 and July 4, 1957, the appearance strongly suggested the same mode of development. There are no incontrovertible examples of material, initially bright, becoming dark, though we suspect that this was so in a number of events, including those of August 24, 1956 and January 18, 1957.

All the evidence is in favour of the initial bright expansion disappearing by becoming transparent, and the dark surge appearing by a transparent stream becoming opaque. We consider this to be the normal behaviour.

Our observations have been made with a 0.7 \AA filter centred on the normal $\text{H}\alpha$ line. The surge $\text{H}\alpha$ line in general suffers Doppler displacements, so that the

THE FLARE-SURGE EVENT

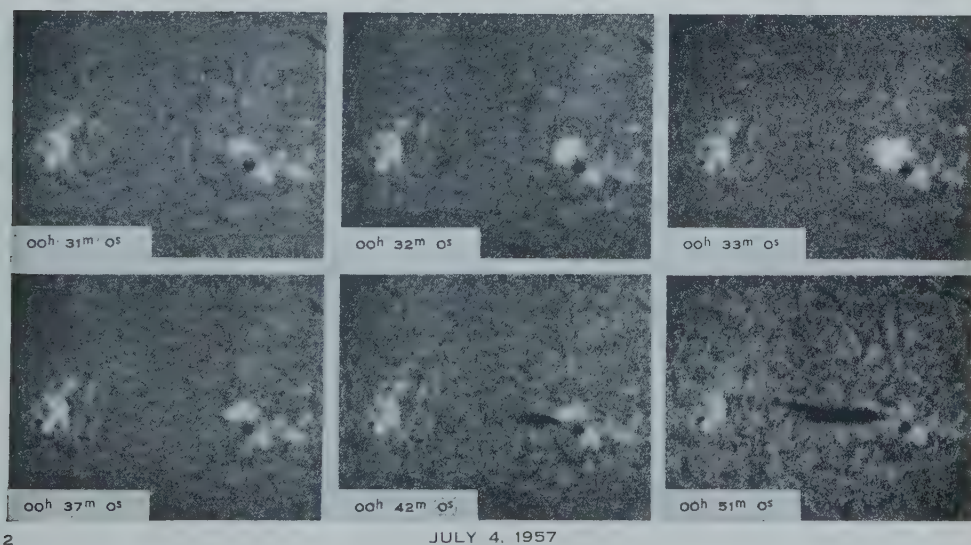
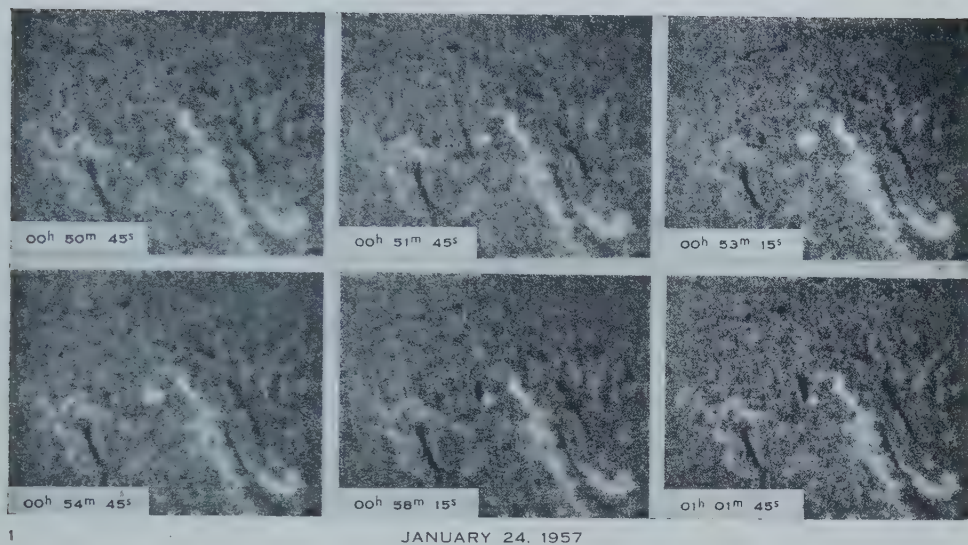


Fig. 1.—Flare-surge on January 24, 1957 (26°N. , 0°W.). No sunspots are visible in these prints, but a very small group was indicated near this position on the Fraunhofer Institut chart for this day.

In this and all subsequent plates the Sun's polar axis runs from top (N.) to bottom (S.), with the east and west limbs respectively on the left and right-hand sides.

Fig. 2.—Flare-surge on July 4, 1957 (13°N. , 23°E.). Note the initial diffuse bright expansion of the flare and the subsequent development of the dark surge almost radially away from the sunspot, rather than from the flare.

THE FLARE-SURGE EVENT

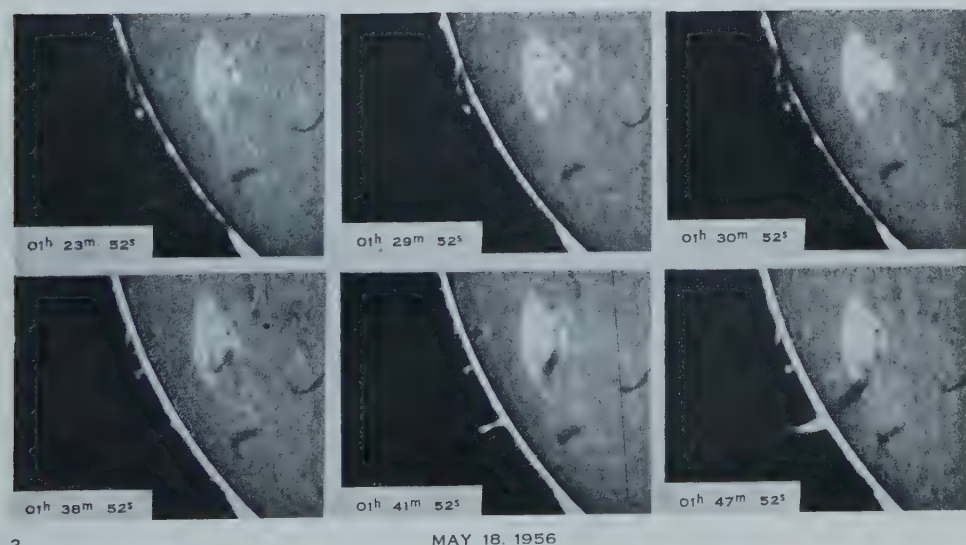
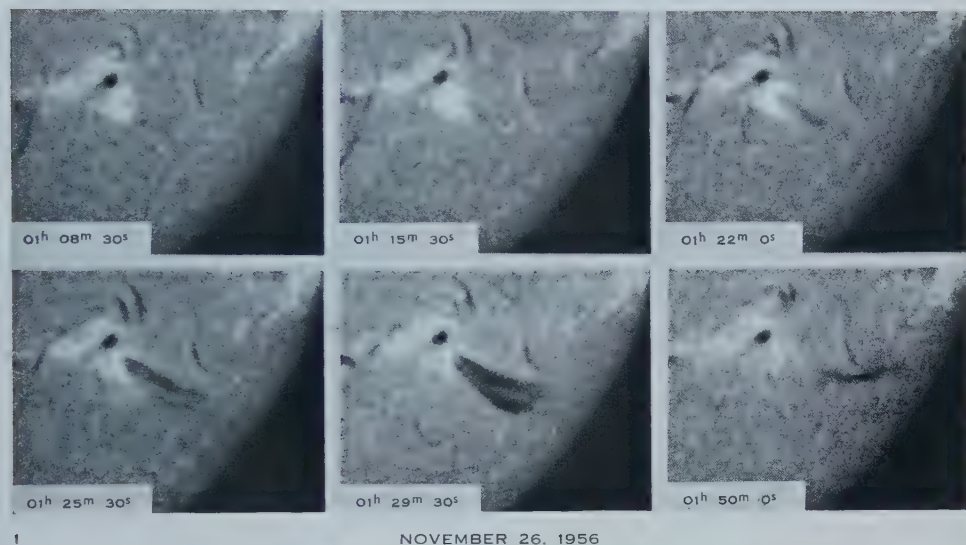


Fig. 1.—Flare-surge on November 26, 1956 (24 °S., 33 °W.). This is one of many such events observed near this spot group. The limb appearance of a flare-surge from the same group is shown in Plate 5, Figure 2.

Fig. 2.—Flare-surge on May 18, 1956 (26 °S., 52 °E.). Enhanced exposures have been given to the limb by covering the image of the Sun's disk after a suitable exposure; faint prominences, otherwise invisible, are then recorded on the same frame as the disk.

The present sequence demonstrates the continuity of emission of the surge particle stream, initially in the form of a bright expansion of the flare, then becoming faint and transparent while moving outwards to be visible beyond the limb on enhanced exposures. The dark opaque surge forms a lower part of the emitted particle stream.

THE FLARE-SURGE EVENT

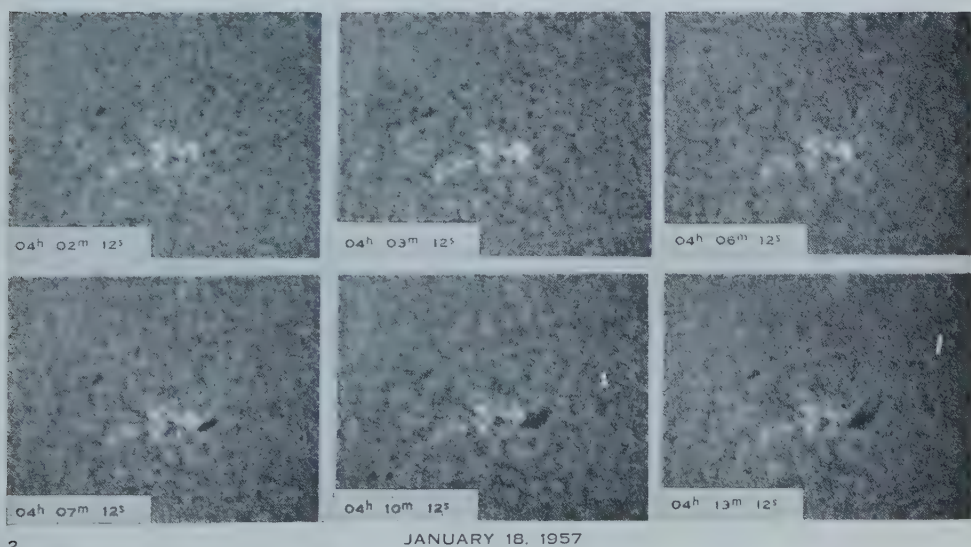
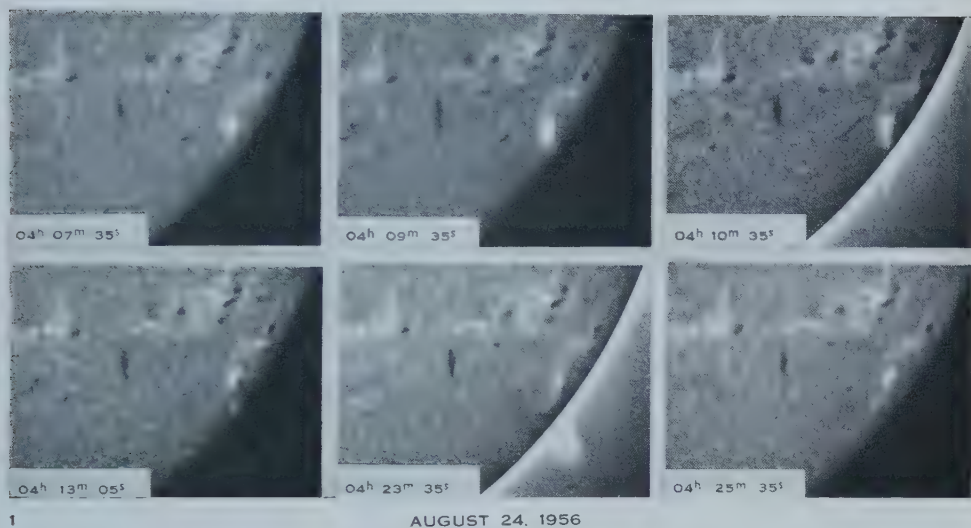


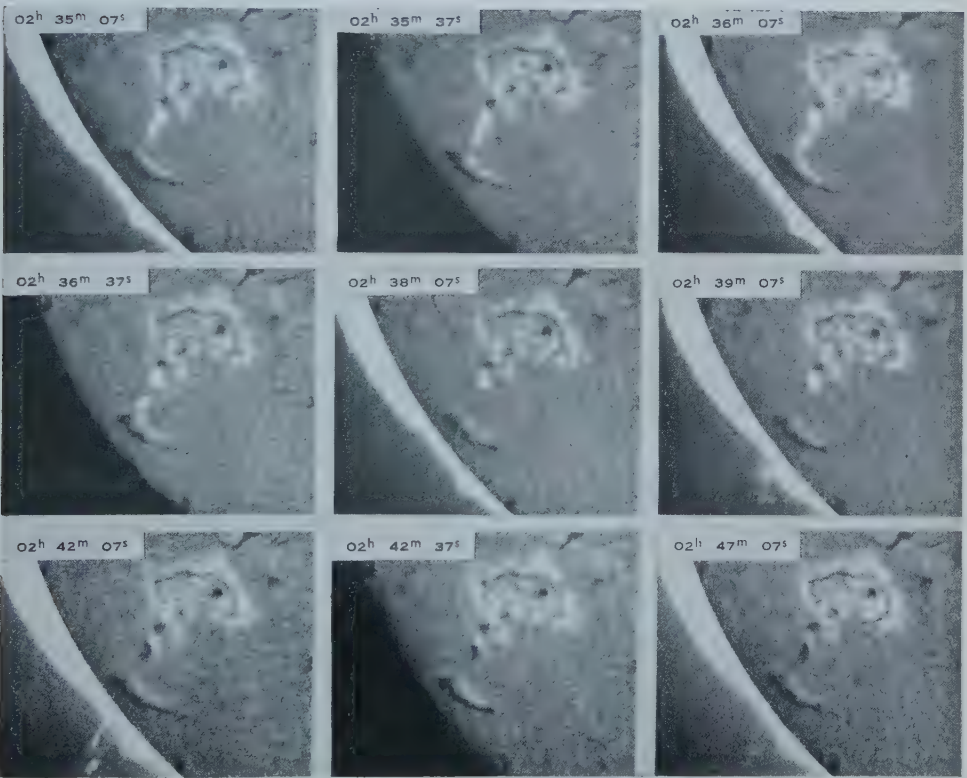
Fig. 1.—Flare-surge on August 24, 1956 (25°S. , 63°W.). Note the diffuse bright expansion of the flare. Enhanced limb photographs at $04^{\text{h}} 10^{\text{m}} 35^{\text{s}}$ and $04^{\text{h}} 23^{\text{m}} 35^{\text{s}}$ show the development of a faint transparent surge prominence during this period.

Fig. 2.—Flare-surge on January 18, 1957 (7°S. , 18°W.). Note the rapid development of the dark surge between $04^{\text{h}} 06^{\text{m}} 12^{\text{s}}$ and $04^{\text{h}} 07^{\text{m}} 12^{\text{s}}$, after which the dark surge has an approximately uniform outward velocity.

THE FLARE-SURGE EVENT



MAY 30, 1956

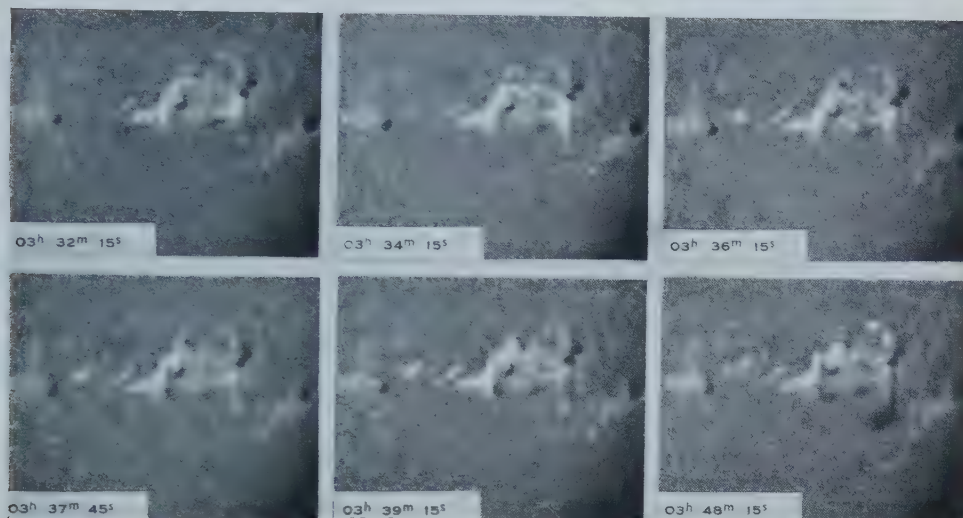


MAY 7, 1957

Fig. 1.—Bright streamer on May 30, 1956 (20°S., 70°W.). An enhanced limb exposure shows the development of the streamer into a faint transparent surge. The tiny flare at the base of the bright streamer lasted from 00h 36m to 00h 43m, and can be seen best by studying the film in cinematographic projection; it is difficult to detect on these prints.

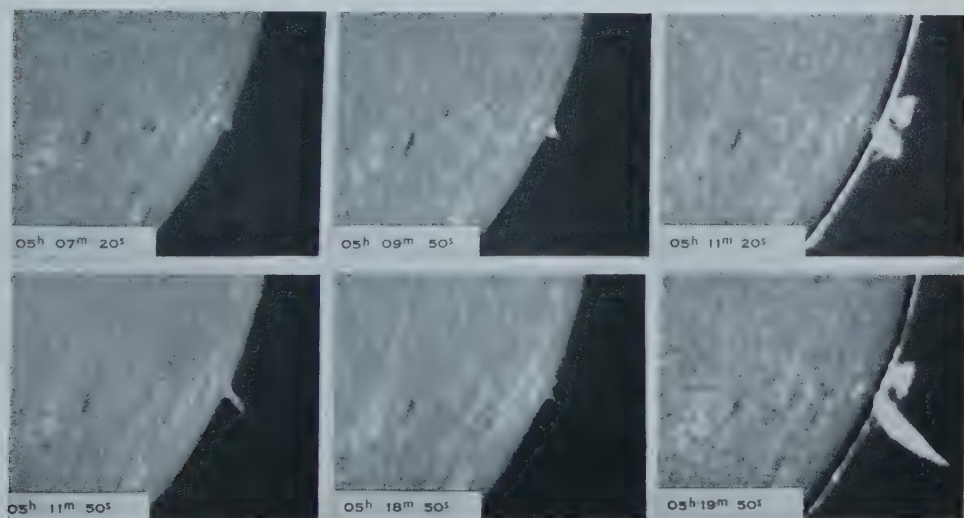
Fig. 2.—Ejected flare on May 7, 1957 (30°S., 55°E.). Enhanced limb exposures show the ejection, now faint and transparent, crossing the limb at 02h 39m 07s; simultaneously a dark surge starts moving outwards along the same path.

THE FLARE-SURGE EVENT



1

AUGUST 23, 1956



2

NOVEMBER 30, 1956

Fig. 1.—Flare-surge on August 23, 1956 (19°S. , 41°W.).

Fig. 2.—Flare-surge at the limb on November 30, 1956 (25°S. , 90°W.). This occurred in the same spot group as the event of Plate 2, Figure 1. Enhanced limb exposures demonstrate the extent of diffuse faint matter surrounding the brighter regions. Maximum flare brightness occurred early in the event.

observed changes in transparency refer to wavelengths displaced by unknown amounts from the surge line centre. There is little doubt, however, that similar changes occur at the centre of the surge $H\alpha$ line.

VIII. EJECTED FLARES

These events, comprising the ejection of November 12, 1956 described by Bray *et al.* (1957) and that of May 7, 1957, are very rare. So far we have not observed such flares on the limb, though Dodson and McMath's observation mentioned in Section I may have related to one of this type.

Both events showed the ejection of part of a flare, the main bulk of the flare remaining stationary. In each case the ejection faded, eventually crossing the limb to appear as a surge. The brightness of the November 12, 1956 ejection passed through a maximum in flight; it did not disappear against the disk, though its much greater extent on the limb, revealed by enhanced exposures, showed much of the material to be highly transparent; it is interesting to note that on falling back towards the Sun's surface some of the matter appeared dark, some slightly bright, against the disk. The ejection of May 7, 1957 became transparent and disappeared before crossing the limb where it, too, was revealed on longer exposures. In each case the bright ejection was followed, some time later, by a small dark surge on the disk.

The events of August 23 and 24, 1956 provided intermediate cases in which bright matter was ejected from and eventually separated from the flare. While brighter than the normal chromosphere, neither ejection was bright enough to be classed as an ejected flare.

IX. DISCUSSION

We have shown that the flare-surge is a common event and that probably all dark surges originate in flares. Matter is expelled from the flare itself or from its immediate neighbourhood, and in most cases the surge appears in contact with the flare. Initially the surge moves in a direction generally radially away from the nearest large sunspot. However, we have been unable to find any preferred location for this type of flare within the spot group, or any spot group characteristic favourable to surge production, though some are obviously far more prolific than others. We have not found any substantial difference in surge productivity between large flares and small.

The essential feature of all these events is the ejection of a stream of particles soon after the beginning of the flare. The first indication is an expansion of the flare, usually asymmetric and diffuse though at times in the form of a streamer, which soon fades and becomes transparent. The particle stream continues on its path, outward though invisible. Usually, though not always, part of the stream becomes opaque above or near the flare and extends outwards as the dark surge.

The varying appearances of the stream depend on changes in the physical conditions. Jefferies (1955, 1956, 1957) has recently shown that both in bright flares and in stable prominences the temperature is of the order of $10-15 \times 10^3$ °K, but that the electron concentration is somewhat higher in flares ($5 \times 10^{11} - 10^{13}$ cm⁻³).

than in prominences (10^{10} – 5×10^{10} cm⁻³). That flares are about as bright in H α as a 6000 °K black body, despite their much higher temperature, but prominences are much fainter is due to their large departures from thermodynamic equilibrium. It follows also that relative temperatures in various parts of a flare-surge cannot be deduced solely from brightness observations.

In this regard it is important to note that in the outward particle stream there is an expansion, demonstrated by the initial bright phase of the flare and by the shape of the dark surges which expand with height. This is also to be expected as the material rises into the low pressure corona. But, at temperatures of the order of 10^4 °K, if the expansion were to occur isothermally or with increasing temperature, the recombination rate would drop rapidly and so would the optical thickness. The increase in opacity during the development of a dark surge therefore indicates that the temperature is decreasing. On the other hand, in the ejected flare described by Bray *et al.* (1957) the brightness increased during flight; according to Jefferies' (1957) analysis, this can happen during expansion only if the temperature increases suitably.

In our view, dark surges, bright surges on the disk which fade without turning dark, and ejected flares are basically similar phenomena, in which a particle stream is expelled from the flare. They differ in rate of heating, density, thickness, and rate of expansion; the one stream can exhibit a wide variety of appearances at various stages of its life.

X. REFERENCES

- BRAY, R. J., LOUGHHEAD, R. E., BURGESS, V. R., and MCCABE, MARIE K. (1957).—*Aust. J. Phys.* **10**: 319.
- BRUZEK, A. (1951).—*Z. Astrophys.* **28**: 277.
- DODSON, HELEN W., and HEDEMAN, RUTH E. (1949).—*Astrophys. J.* **110**: 242.
- DODSON, HELEN W., and HEDEMAN, RUTH E. (1952).—*Observatory* **72**: 30.
- DODSON, HELEN W., and McMATH, R. R. (1952).—*Astrophys. J.* **115**: 78.
- ELLISON, M. A. (1942).—*Mon. Not. R. Astr. Soc.* **102**: 11.
- GIOVANELLI, R. G. (1940).—*Astrophys. J.* **91**: 334.
- JEFFERIES, J. T. (1955).—*Mon. Not. R. Astr. Soc.* **115**: 617.
- JEFFERIES, J. T. (1956).—*Mon. Not. R. Astr. Soc.* **116**: 629.
- JEFFERIES, J. T. (1957).—*Mon. Not. R. Astr. Soc.* **117**: 493.
- NEWTON, H. W. (1942).—*Mon. Not. R. Astr. Soc.* **102**: 2.
- SEVERNY, A. B., and SHAPOSHNIKOVA, E. F. (1954).—*Astr. J., Moscow* **31**: 124.

POLARIZATION MEASUREMENTS OF THE THREE SPECTRAL TYPES OF SOLAR RADIO BURST

By M. KOMESAROFF*

[*Manuscript received December 19, 1957*]

Summary

A swept-frequency technique was used for measuring the polarization of solar radio bursts occurring at the beginning of the present sunspot cycle. Of special interest were the results for bursts of spectral type III. Contrary to the inference drawn from earlier work, it was found that many of these bursts are highly polarized. Furthermore, there were strong indications that the polarization is produced at the radiation source and is not imposed by propagation conditions in the overlying media.

I. INTRODUCTION

Radio spectroscopic observations of intense solar disturbances have shown that these emissions may be classified into a number of distinct spectral types (Wild and McCready 1950). Observations at discrete frequencies had first suggested the possibility of such a classification, and they had also indicated that each type of burst has its own polarization characteristics (Payne-Scott 1949; Payne-Scott and Little 1951).

During "noise storms" (spectral type I) the polarization was found to be often almost complete, while outside noise-storm periods randomly polarized bursts were found to occur. However, Payne-Scott and Little also recognized certain large "outbursts" whose polarization was random during their initial phase and later became elliptical. The identification of Payne-Scott's "unpolarized bursts" with spectral types is not clear. Many of these were certainly of spectral type III, and it was suggested by Wild and McCready that the type-III burst characteristically showed no circular polarization.

In the investigation reported here, the spectrum and polarization of solar disturbances have been measured simultaneously over a wide range of frequencies. The observations were made using a modification of the spectroscope described previously (Wild, Murray, and Rowe 1954) and the technique (details of which are given in Section II) provides simultaneous records of the dynamic spectrum and polarization. Thus it has the advantage over previous methods of measuring polarization, that it provides a positive identification of the spectral type under investigation. In addition measurements can be made rapidly, permitting the study of short-lived spectral features. A broad survey is obtained of polarization as a function of time and frequency rather than accurate measures of the polarization ellipse at individual frequencies. The results largely confirm the suggested association between polarization and spectral type, but they also reveal a new

* Division of Radiophysics, C.S.I.R.O., University Grounds, Chippendale, N.S.W.

feature. A considerable number of type III bursts have been found to be strongly polarized. A few of the polarized bursts displayed a "harmonic" structure (see Section III (c)) and, in the cases examined, the sense of rotation of the harmonic component agreed with that of the fundamental. The degree of polarization of the two components appeared to differ, however.

A preliminary report on this work was given by Wild (1955) at the I.A.U. symposium on radio astronomy at Jodrell Bank.

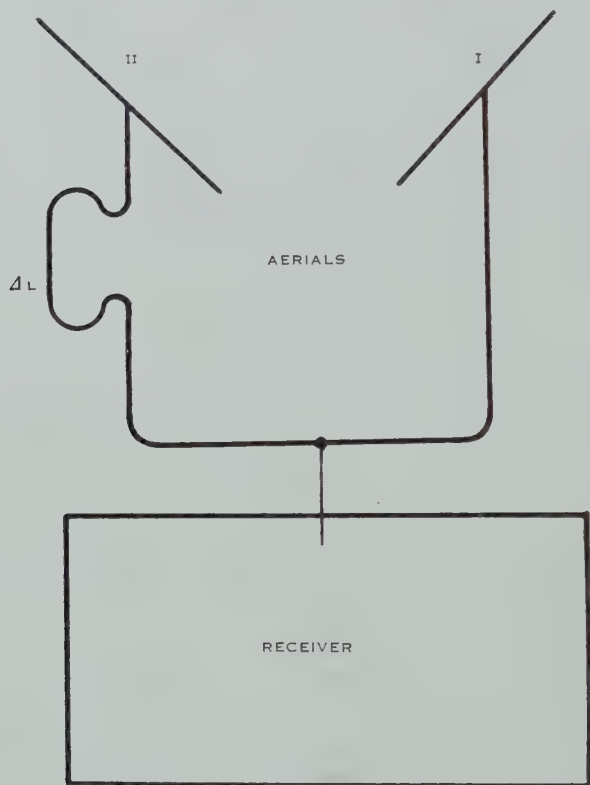


Fig. 1.—Block diagram of the equipment for measuring polarization. One unit of this type was used for each of the three frequency ranges which together covered the band 40–240 Mc/s.

II. A SWEEP-FREQUENCY TECHNIQUE FOR MEASURING POLARIZATION

(a) *The Method*

The method of polarization measurement is a variant of the crossed-aerial technique, adapted to a swept-frequency instrument.

Three pairs of mutually perpendicular rhombic aerials, each pair equatorially mounted and mechanically driven to follow the Sun, are used to cover three adjacent bands of the 40–240 Mc/s frequency range. Each aerial pair has a corresponding receiver and the individual aerials of each pair are connected to the common receiver by twin-wire transmission lines of different lengths, as shown in Figure 1. If ΔL is the difference in line length, the system accepts one

circularly polarized component when the receiver is tuned to a wavelength λ such that $\Delta L = (2n \pm \frac{1}{2})\lambda$, where n is an integer, the sign being determined by the sense of rotation of the electric vector. As the receiver tuning is varied the output goes through a series of maxima and minima. Examples of such patterns for type I radiation are shown in the record of Plate 1. (It should be noticed that the frequency scales are non-linear and that different values of ΔL apply to the different frequency ranges.)

In the general case of partial elliptical polarization, the plane-polarized components may be unequal. The phase angle θ , by which the plane-polarized wave component parallel to aerial II lags that parallel to aerial I, may assume any value between $\pm\pi$ radians. If the type of polarization remains constant across the frequency range and the receiver is continuously tuned, a record is obtained of the radiation power spectrum modulated by a sinusoidal pattern whose maxima occur at wavelengths λ_n such that $(\theta + 2\pi\Delta L/\lambda_n) = 2n\pi$, where n is an integer. Hence θ may be determined from the frequencies at which these maxima occur, since

$$\theta = 2\pi(n - f_n/\Delta f), \quad \text{where } \Delta f = c/\Delta L \text{ and } f_n = c/\lambda_n,$$

c being the velocity of electromagnetic radiation. Figure 2 is a schematic representation of a record for the case where $0 < \theta < \pi$.

In order to specify the polarization completely at any one frequency, we need four independent measures of the radiation parameters. However, if we assume that the polarization does not change rapidly with frequency, we may derive considerable information from the modulation pattern alone.

The sense of rotation may be immediately determined from the phase angle θ , being left-handed for $0 < \theta < \pi$ and right-handed for $-\pi < \theta < 0$. (The convention adopted by the International Astronomical Union is followed in describing sense of rotation. The polarization is right-handed if the vector, viewed along the direction of propagation of the ray, rotates in a clockwise sense.)

The percentage polarization may be defined as

$$100(E_1^2 + E_2^2)/(2R^2 + E_1^2 + E_2^2),$$

where E_1 and E_2 are the phase-coherent voltages received on the individual aerials and R is the random component of voltage received on each aerial.

For circular polarization the percentage polarization is equal to the percentage modulation of the power spectrum given by

$$100(P_{\max.} - P_{\min.})/(P_{\max.} + P_{\min.}) \quad (\text{see Fig. 2}).$$

This is approximately true also for fairly broad ellipses, but, as the modulation for a given percentage polarization usually decreases as the ellipse becomes elongated, a knowledge of the degree of modulation enables us only to set a lower limit to the percentage polarization.

It is possible to estimate the shape of the polarization ellipse, which may be specified by a parameter p , defined as the ratio of the minor to the major axis

of the ellipse. Sufficient information is not available to determine this parameter from individual measurements at single frequencies, but a mean value may be calculated for activity extending over a range of frequencies.

It is known that, when polarized radio waves traverse an ionized medium in the presence of a magnetic field, the polarization ellipse undergoes rotation due to the Faraday effect, the amount of the rotation being a function of both frequency and path length in the medium (Cohen 1956). Murray and Hargreaves (1954) have shown that even at 100 Mc/s there may be several complete rotations in the ionosphere, and a much larger effect may be expected in the solar corona. The magnitude of this effect is sufficiently great for polarization ellipses observed at the different frequencies within the range of our equipment to show all possible orientations. Now an ellipse whose axes lie in the aerial planes yields a phase

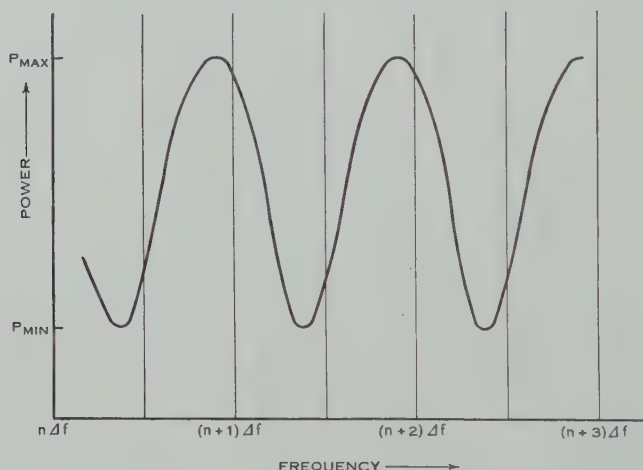


Fig. 2.—Schematic representation of a record of polarized radiation having a flat spectrum. The frequency interval Δf is defined in Section II (c).

measurement of $+\frac{1}{2}\pi$ or $-\frac{1}{2}\pi$, and, as the axes rotate from this position about the direction of propagation, $|\theta|$ departs from $\frac{1}{2}\pi$ at a rate which depends on the parameter p . Therefore the probability of observing a particular value of θ is determined by the shape of the ellipse. The problem is formulated in Appendix I, and probability histograms for various values of p are presented in Figure 3 (a). On the assumption that the ellipse shape does not vary rapidly with either frequency or time, we may compare the results of a large number of phase measurements extended across a range of frequencies with this set of computed histograms and so obtain a mean value of p .

(b) Calibration

The intensity scale is calibrated by injecting signals from a diode noise generator into the receiver, and the frequency scale is calibrated against the harmonics of a crystal oscillator.

The differences in length ΔL of the pairs of feeders are measured by a standing wave technique. Each line is short-circuited at the aerial and fed by a signal generator at the receiver end. The approximate length being known, the exact length is determined from the positions of standing wave minima near the receiver, the determination being carried out at a number of frequencies. As in each case it is required to determine the difference of two lengths and not their individual values, no error is introduced by imperfections in the short circuits.

(c) Accuracy

The two quantities which are measured from the records are the phase angle and percentage modulation. Because many of the measurements, particularly in the case of type III bursts, had to be made from intensity-modulated records the "probable errors" quoted refer to this type of display. Measurements on "A scan" records yielded somewhat better accuracy. (For a description of the two methods of display see Wild, Murray, and Rowe (1954).)

(i) *Measurement of Phase Angle*.—Reading error is the most serious limitation to the accuracy with which the phase angle θ may be determined. The error is a function of both intensity of activity and frequency and its assessment is difficult. The mean "probable error" of phase measurement across the low and medium frequency ranges is estimated to be about $\pm \frac{1}{8}\pi$ ("probable error" is taken to mean that scatter which includes 50 per cent. of observations).

(ii) *Measurements of Percentage Modulation*.—The probable error in the measurement of percentage modulation is also a function of intensity of activity and is estimated to be about 25 per cent. It is unlikely that the modulation would be detected if its absolute value were less than this.

III. THE POLARIZATION OF THE THREE SPECTRAL TYPES

Polarization measurements summarized here began in January 1955—the beginning of the new sunspot cycle. The data for type I and type III bursts cover the period between January and October 1955, and the type II burst data cover most of 1955 and 1956. The new feature revealed by the measurements is that a substantial number but by no means all of the type III bursts are strongly polarized. From earlier observations it had been inferred that their polarization was random.

Because of calibration and other difficulties with the high frequency range of the equipment, most of the results presented here were derived from the low and medium frequency records covering the band 40–140 Mc/s.

(a) Type I Bursts (*Storm Bursts*)

The records of 13 days' activity were examined, and the conclusion of earlier observers confirmed, that Type I bursts display a high degree of quasi-circular polarization. The radiation was polarized at all frequencies and usually showed the same sense of rotation throughout the frequency range and during any one day. For the more intense storms during this period the sense of rotation was left-handed. The sense of rotation data are presented in Table 1 (a). It will be seen that there was only one case in which the sense of rotation changed

throughout the day—this was on January 10, 1955, when there was a sporadic storm during which both senses were observed, but on different frequency ranges and at different times. The letters L.F. and M.F. in the table refer to the low (40–70 Mc/s) and medium frequency (75–140 Mc/s) ranges of the equipment respectively.

Generally the polarization was greater than 50 per cent.

TABLE I
SENSE OF ROTATION OF POLARIZATION ELLIPSE

(a) Type I Bursts			(b) Type III Bursts		
Date (1955)	U.T.	Sense of Rotation	Date (1955)	U.T.	Sense of Rotation
Jan. 6	0423–0845	L			
Jan. 10	0434–0450	L (M.F. range only)	Jan. 10	0801	L
	0514–0521	R (L.F. range only)		2052–2054	L
Jan. 11	0338–0530	L			
Jan. 12	0834–0837	L ?	Jan. 12	0705–0815	L
				2140	L
Jan. 13	2318–0634	L			
Jan. 15	0351–0841	L			
Jan. 16	2216–0511	L			
	2150–2211	L			
Feb. 2	0226–0436	R	Feb. 2	0634	R
				0712	R
				0754	R
			Mar. 3	0300	R
Mar. 4	0435–0488	R			
Aug. 12	0307–0609	R	Aug. 12	0536	R
				0538	R
Sept. 6	0003–0720	L			
Oct. 10	2242				
Oct. 11	0714	L			

Using the method outlined in Section II and Appendix I, frequency distributions of phase measurements have been prepared for some of the longer-duration noise storms, and are presented as histograms in Figure 3 (b), together with the set of histograms computed for a number of values of p in Figure 3 (a). It would appear from this figure that p generally lay between 0.5 and 1.0.

(b) Type II Bursts

Polarization observations were carried out for 13 type II bursts. Eight of these records were completely unmodulated. In the case of four of the remainder no conclusion was reached. For two of these, very weak patterns were observed which may have been spectral fluctuation. The other two occupied so narrow a band of frequencies that the observation of a modulation pattern would probably have been impossible.

Only one record was probably polarized for a short period but the modulation at its maximum was not greater than 30 per cent.

(c) Type III Bursts

In the course of the present investigation about 500 type III spectra were recorded, and of these, some 50 per cent. exhibited the modulation pattern characteristic of fairly strong polarization (examples of polarized and unpolarized type III records are shown in Plate 2).

It is possible that terrestrial effects were responsible for some of the modulation. This was first suggested by the observation that, in the period January–May 1955, the ratio of modulated to unmodulated bursts rose sharply for the higher

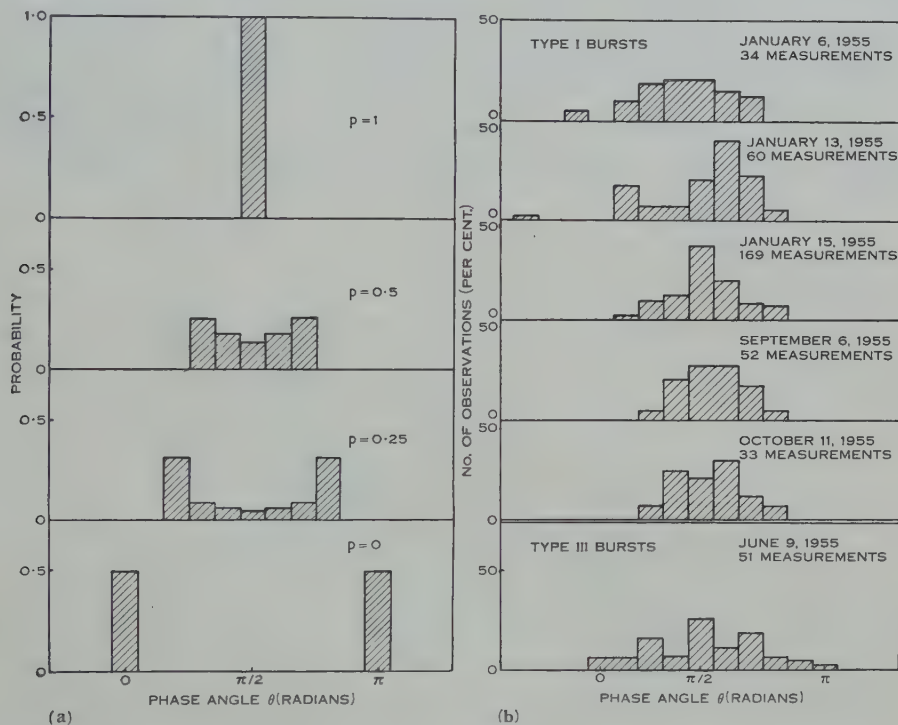


Fig. 3 (a).—Probability distributions of phase angle θ , for various values of the parameter p , calculated according to the method described in Appendix I.

Fig. 3 (b).—Observed frequency distributions of phase angle for type I and type III radiation occurring on the days indicated.

solar zenith angles. Two terrestrial mechanisms capable of producing spurious polarization effects are differential absorption of oppositely polarized modes in the ionosphere and ground reflection.

It can be shown that polarization produced by differential absorption in the ionosphere is not likely to exceed 5 per cent. Moreover, Slee (unpublished observations), observing at 85 Mc/s, found that the polarization of the source in Taurus (at about 56° from zenith) was less than 2 per cent. These figures are below the threshold of detectability of our equipment.

The effect of ground reflection is more difficult to assess. The broad lobes of the rhombic aerials are capable of receiving both direct and ground-reflected

radiation components over a large range of solar zenith angles, and these components may give rise to interference patterns. Determination of the amplitudes of these patterns involves extremely lengthy calculation and has been carried out for only a few special cases. However, the fact that only three of a total of thirteen type II records show evidence of (extremely faint) modulation, suggests that over considerable periods the effect is not large.

Although terrestrial effects cannot be completely ruled out, there is considerable evidence of a solar origin for much of the type III polarization. This may be summarized as follows:

(1) Observations for the period May–October 1955 do not show the trend with zenith angle shown by the earlier results.

(2) Measurements of phase angle have been made for about 40 bursts and are presented in the histogram of Figure 4. Clustering of phase values about $+\frac{1}{2}\pi$ and $-\frac{1}{2}\pi$ indicates components of genuine quasi-circular polarization.

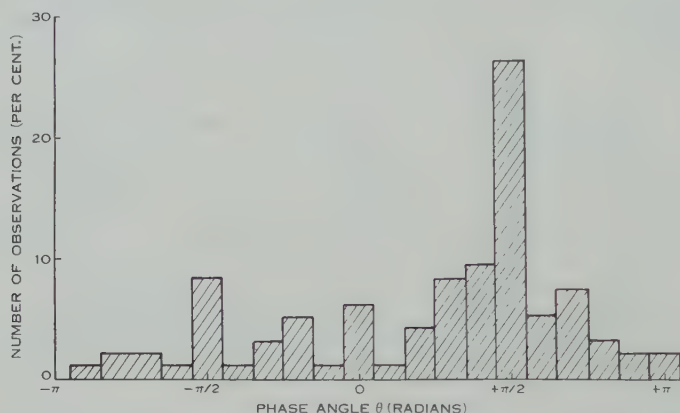


Fig. 4.—Frequency distributions of phase angle of 40 type III bursts occurring throughout the year.

(3) Polarization of type III bursts occurring on the same day showed the same sense of rotation; furthermore, this was the same as that of any type I radiation occurring within a period of one or two days (see Table 1).

(4) The occurrence of bursts showing a harmonic structure (see below) in which the fundamental component is more deeply modulated than the harmonic, cannot be explained in terms of any local mechanism.

To establish more definitely the solar origin of the polarization of type III bursts and to determine the features of this polarization, a sample record was chosen for detailed analysis, the record being one for which instrumental effects appeared to be negligible. The record selected was that of June 9, 1955. Table 2 is a summary of this day's activity.

Both highly polarized and unpolarized bursts were observed between 0008 and 0248 U.T. during which period the solar zenith angle changed by only about 6.5° . Furthermore, polarized bursts were observed at 0129 and 0153 U.T. on the previous day, June 8, whereas bursts occurring at 0118 on June 9 were

unpolarized. The depth of modulation to be expected from ground reflection was calculated for these bursts and no significant difference was found between the depth of modulation due to ground effect at the times when the records were modulated and those when they were not. In any case the effect was near the threshold of detectability of the equipment.

As there were no "A scan" records available, a microphotometer was used to evaluate the depth of modulation of the cluster of bursts occurring between 0246 and 0250 U.T. This was found to vary between about 30 and 70 per cent. Phase measurements made on this same cluster are presented in the histogram of Figure 3. Although there is considerable "scatter", it can be seen that the values cluster around $\frac{1}{2}\pi$, indicating that the polarization was elliptical, with the left-handed sense of rotation.

TABLE 2
SUMMARY OF ACTIVITY JUNE 9, 1955

Time. (U.T.)	Bursts	Observed Modulation
0001-0002	Type III	None
0008-0019	Main part of type II burst	None
0032	Tail of type II burst	No evidence of modulation
0033 $\frac{1}{2}$	Cluster of type III bursts	High
0118-0120	Cluster of type III bursts	None
0246 $\frac{1}{2}$ -0250	Cluster of type III bursts some showing harmonic structure	High—up to 70 per cent.

Several "harmonic" bursts were observed during the present investigation, and of these some were polarized, though not all. The spectrum of these bursts consists of two components, one duplicating the features of the other at about twice the frequency. Their occurrence was first reported by Wild, Murray, and Rowe (1954), who attributed the spectral structure to the emission of a fundamental frequency and its second harmonic from the same coronal level.

Of the cluster of (polarized) bursts observed between 0246 and 0250 U.T. on June 9, two show a well-defined harmonic structure. This record is reproduced in Plate 2. In each case it appears that the fundamental is more deeply modulated than the harmonic. This visual impression is confirmed by microphotometer measurements, and Figure 5 is a plot of degree of modulation versus frequency for each component of these two bursts.

IV. DISCUSSION OF RESULTS

(a) Comparison with Other Observations

As noted earlier, the present work confirms the conclusion of earlier observers (notably Payne-Scott and Little 1951) who have reported that the bursts now known as type I ("storm bursts") are strongly polarized. For bursts of spectral types II and III, a perfectly definite comparison of the present results with those

of earlier observers is not possible because of the uncertainties involved in attempting to identify features observed at single frequencies with those observed with the spectrometer.

Our observations of type II bursts have indicated that polarization, if it is present at all, is very slight. Payne-Scott and Little (1951), on the other hand, reported surges of activity whose intensities and durations would suggest type II bursts, but they found that, although these were initially unpolarized, they often went through a period of elliptical polarization. The apparent conflict may be resolved if we assume that the period of elliptical polarization represented the beginning of a type I storm. The single-frequency records do not, however, provide sufficient information to settle this point.

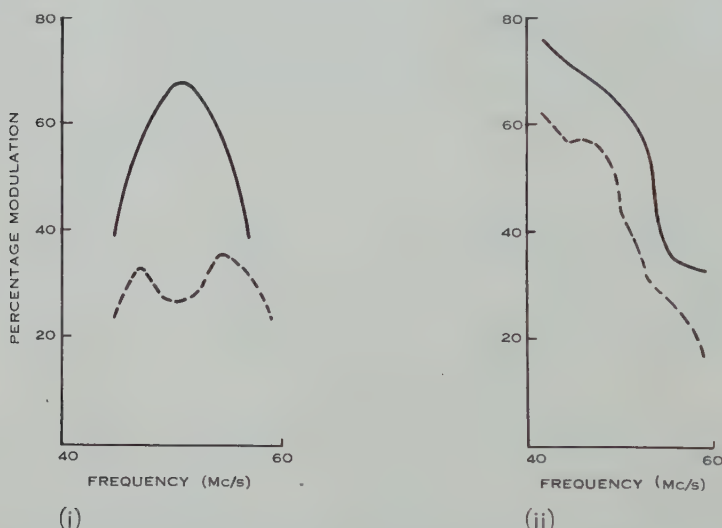


Fig. 5.—Degree of modulation as a function of frequency for the two polarized harmonic type III records of June 9, 1955 shown in Plate 2. The full line refers to the fundamental and the broken line to the harmonic component in each case.

Type III bursts, largely as a result of Payne-Scott's earlier work (1949), have been thought to be unpolarized, whereas it is clear from the present investigation that many, though not all, are strongly polarized. In this case the conflict would appear to be terminological rather than real. Payne-Scott used "unpolarized burst" as a definitive term, and showed that short-lived bursts of this kind occurred almost simultaneously at widely separated frequencies, that is, that they corresponded to what were later designated spectral type III. She did not discuss the converse problem—whether all bursts having this spectral feature were unpolarized. Single-frequency records available in this laboratory have been examined in the light of the present results, and a number of examples have been found in which bursts occurring nearly simultaneously at 85 and 60 Mc/s were quite highly polarized. In the case of type III bursts there is, therefore, no conflict between the earlier and the present results.

(b) *The Origin of Polarized Radiation*

The present results enable us to draw some general conclusions about the origins of type I and type III polarization. We may explain the occurrence of polarized radiation either in terms of the generation mechanism or in terms of propagation conditions in the medium between the source and the observer. Polarization imposed by the medium may arise in two ways, as discussed below.

(i) *Polarization by Total Internal Reflection in a Birefringent Medium.*—

The radiation may pass through a region in which the refractive index for one propagation mode is zero. In this case only the mode for which the refractive index remains finite will be transmitted, and the radiation will be completely polarized. Now, it is well known that, for radiation to escape from the corona (see for example Smerd 1950), it must originate above the level for which the refractive index for the ordinary propagation mode is zero. The height of this level is determined by the electron density distribution. In the presence of a magnetic field, there is another level, above the first, at which the refractive index for the extraordinary mode is zero. The separation of these two levels increases with the intensity of the magnetic field. Radiation generated in the region between them is polarized because only the ordinary mode escapes. For quite a wide range of angles between the direction of propagation and the magnetic field, the polarization is very nearly circular. If radiation extending over a broad frequency range is to be polarized in this way, each frequency component must originate between the corresponding levels of zero refractive index. Figure 6 is a plot of the heights of the levels of zero refractive index in an extreme case—above a very large unipolar spot group whose magnetic field at the chromospheric surface is 3600 oersteds. For more common spots the separation between the two levels is less than that shown.

It is possible to explain the polarization of type I radiation in terms of total internal reflection, but such an explanation implies that the radiation source is vastly extended in space. It can be seen from Figure 6 that, for radiation occupying the frequency band 40–140 Mc/s to be polarized by this means, its source must extend over a radial distance of at least 10^5 km.

On the other hand, we cannot explain the polarization of harmonic type III bursts in terms of total internal reflection. The fundamental and harmonic frequency components originate from the same source (Wild, Murray, and Rowe 1954), and, referring again to Figure 6, we see that in no case do the regions between the two reflection levels overlap for frequencies whose ratio is 2 : 1 when the lower frequency lies in the range 20–40 Mc/s.

(ii) *Polarization by Differential Absorption.*—A second way in which the medium may impose polarization is by differential absorption of modes, an effect which depends on the electron collision frequency. Although differential absorption possibly contributes to the polarization of the fundamental, it is quite inadequate to explain the high polarization of the harmonic components of type III bursts. Because the absorption is most marked in the vicinity of the region at which the refractive index is zero, the polarization impressed on the second harmonic radiation (which is propagated outwards from a region

in which the frequency of zero refractive index is only about one-half of the wave frequency) is negligible.

It is difficult to escape the conclusion that at least the type III bursts showing harmonic structure originate in a source which emits polarized radiation.

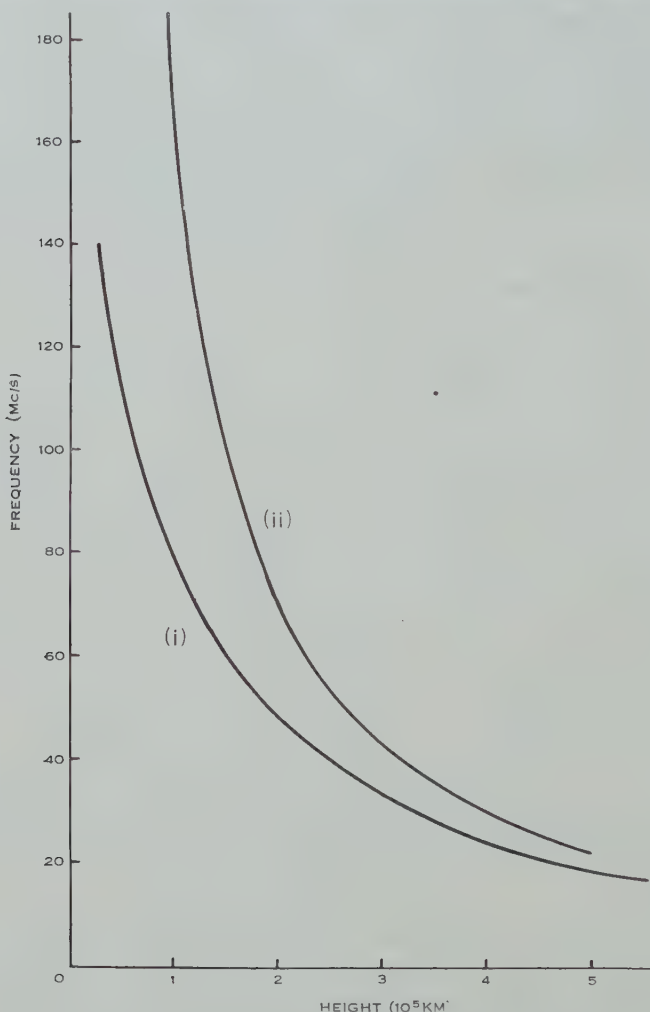


Fig. 6.—Heights of the levels of zero refractive index in the corona above an extremely large unipolar sunspot group having a surface magnetic field intensity of 3600 oersteds. The field and coronal models used are those described by Smerd (1950). The numbers (i) and (ii) refer to the ordinary and extraordinary mode respectively.

V. ACKNOWLEDGMENTS

The author wishes to thank Mr. J. P. Wild, who suggested the present investigation and who, together with Dr. J. A. Roberts and Mr. S. F. Smerd, gave much help in the preparation of this paper.

He also wishes to thank Mr. J. Joice for assistance with observations and measurements.

VI. REFERENCES

- COHEN, M. H. (1956).—Interpretation of radio polarization data in terms of Faraday rotation. Cornell University, Ithaca, N.Y., School of Elec. Eng. Research Reports EE 295.
- MURRAY, W. A. S., and HARGREAVES, J. K. (1954).—*Nature* **173**: 944.
- PAYNE-SCOTT, RUBY (1949).—*Aust. J. Sci. Res. A* **2**: 214.
- PAYNE-SCOTT, RUBY, and LITTLE, A. G. (1951).—*Aust. J. Sci. Res. A* **4**: 508.
- SMERD, S. F. (1950).—*Proc. Instn. Elec. Engrs. III* **97**: 447.
- WILD, J. P. (1955).—Spectral observations of solar activity at metre wavelengths. I.A.U. Symposium, Jodrell Bank, Manchester.
- WILD, J. P., and MCCREADY, L. L. (1950).—*Aust. J. Sci. Res. A* **3**: 387.
- WILD, J. P., MURRAY, J. D., and ROWE, W. C. (1954).—*Aust. J. Phys.* **7**: 439.

APPENDIX I

Estimate of Ellipse Shape based on Many Phase Measurements

Consider elliptically polarized radiation incident on a pair of crossed rhombic aeriels, the direction of propagation being along the common aerial axis.

Then the voltages induced in the aeriels are proportional to the coordinates, with respect to a pair of Cartesian axes lying in the aerial planes, of a point describing an ellipse having the same shape and orientation. Therefore, if the ellipse axes lie in the aerial planes, the voltages are proportional to x_1 and x_2 where

$$\begin{aligned}x_1 &= \cos \omega t, \\x_2 &= p \sin \omega t,\end{aligned}$$

and p is the ratio of the minor to the major axis. If the ellipse rotates through an angle φ about the direction of propagation, then $x_1 x_2$ must be replaced by $x'_1 x'_2$, where

$$\begin{aligned}x'_1 &= x_1 \cos \varphi - x_2 \sin \varphi, \\x'_2 &= x_2 \cos \varphi + x_1 \sin \varphi.\end{aligned}$$

Therefore,

$$\begin{aligned}x'_1 &= (\cos^2 \varphi + p^2 \sin^2 \varphi)^{\frac{1}{2}}. (\cos \omega t + a), \\x'_2 &= (p^2 \cos^2 \varphi + \sin^2 \varphi)^{\frac{1}{2}}. (\cos \omega t + b),\end{aligned}$$

where

$$\begin{aligned}a &= \text{artan } (p \tan \varphi), \\b &= -\text{artan } (p \cot \varphi).\end{aligned}$$

Therefore the angle θ by which x'_2 lags x'_1 is

$$\theta = a - b,$$

and

$$\tan \theta = \frac{p}{1-p^2} 2 \operatorname{cosec} 2\varphi. \quad \dots\dots\dots (1)$$

Now, if an ellipse of given shape (denoted by p) may assume all orientations in the plane perpendicular to the direction of propagation and all orientations are

equally probable, then the probability of observing a particular value of θ , lying in the range $\theta_1 - \theta_2$, is proportional to the corresponding range, $\varphi_1 - \varphi_2$, given by equation (1).

Hence, using equation (1), we may construct a set of histograms, each for a constant value of p , indicating the way in which the probability of observing values of θ lying in a range centred on θ_0 , varies with θ_0 . This has been done in Figure 3 (a).

EXPLANATION OF PLATES 1 AND 2

PLATE 1

Sample records of strongly polarized type I radiation. The bright streaks parallel with the time axis are maxima of the modulation patterns. Unmodulated sections of record obtained with single aeriels are included for comparison in each case. The faint sloping lines crossing the records are due to radiation from power lines having a 50 c/s modulation.

(a) January 13, 1955 at about 0040 U.T.

(b) January 15, 1955 at about 0800 U.T. This record shows the effect of reversing the aerial connection, causing the pattern to shift by one half-fringe width.

PLATE 2

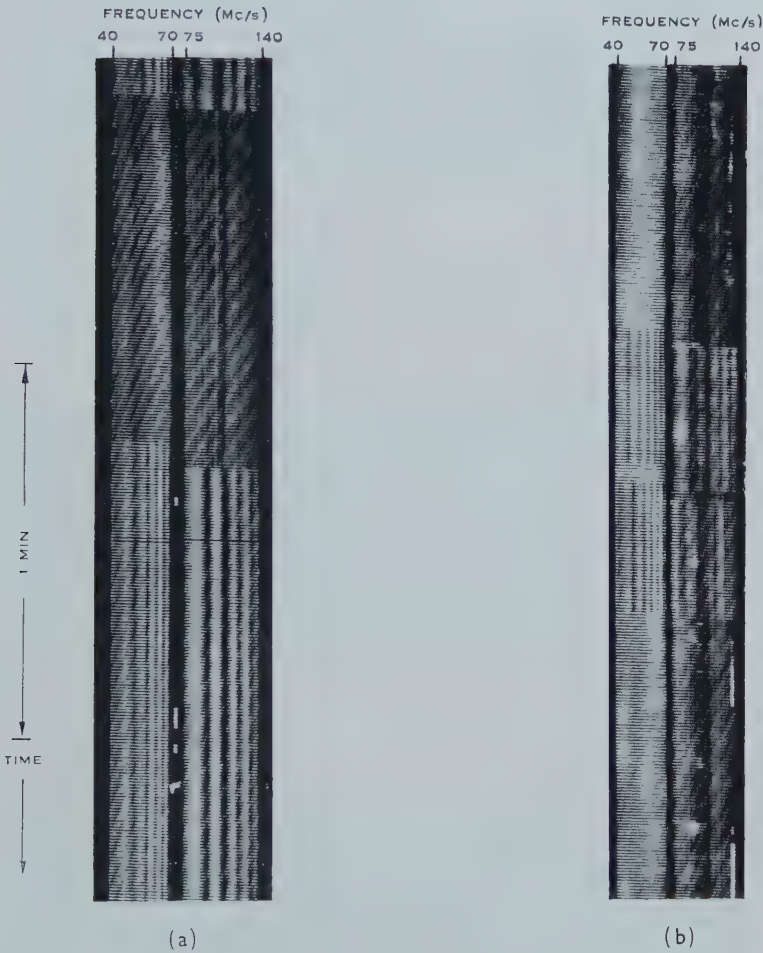
Sample records of type III bursts.

(a) Record of a typical group of unpolarized type III bursts observed on August 15, 1955 at 0155 U.T. The dark spaces on this record and on records (b) and (c) at 70 and 140 Mc/s are produced by switching of the three receivers which together cover the frequency range. The absence of a modulation pattern of vertical dark lines in this record (cf. records (b) and (c)) shows that the radiation was unpolarized.

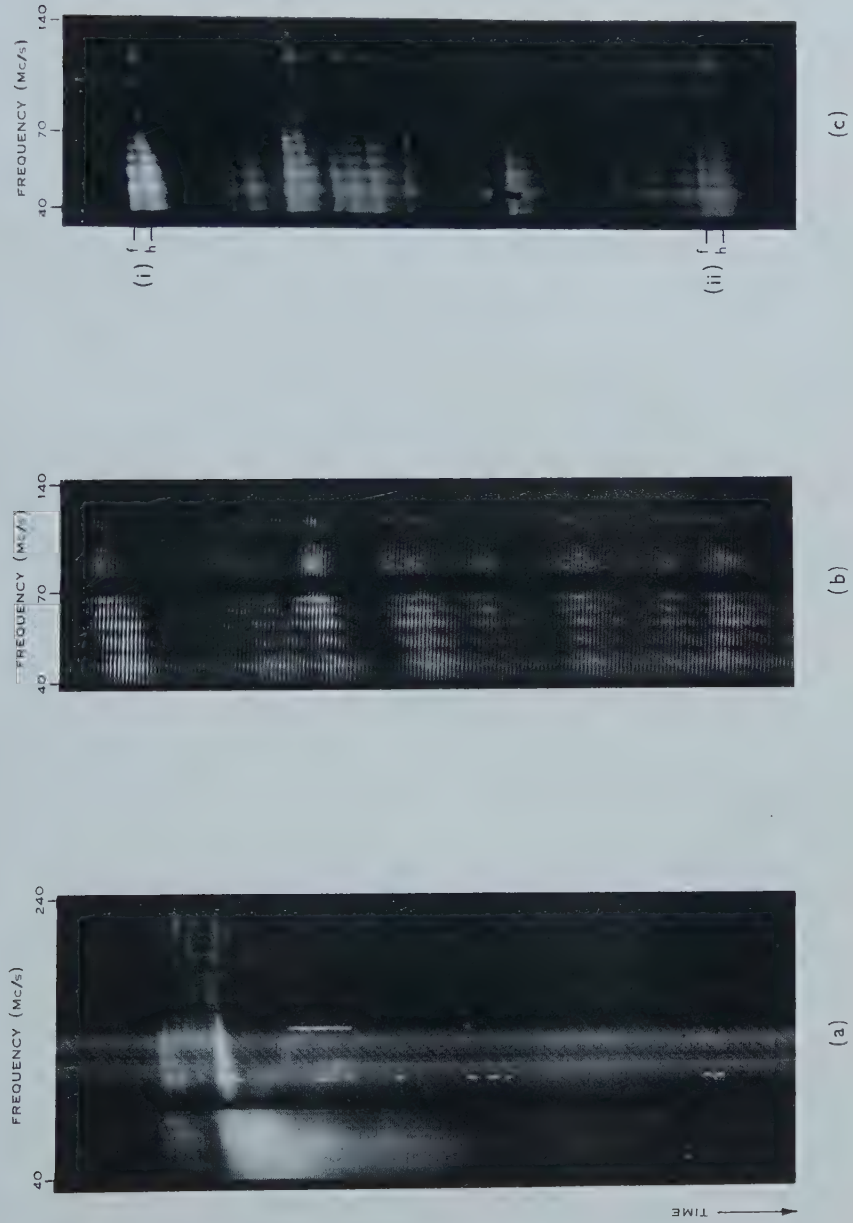
(b) Polarized type III bursts observed on January 12, 1955 at 0749 U.T. These bursts were part of a polarized type III "storm" lasting several hours.

(c) Cluster of polarized type III bursts observed on June 9, 1955 between 0246 and 0250 U.T. The numerals (i) and (ii) indicate bursts showing a definite harmonic structure. The fundamental and harmonic components are denoted by the letters f and h respectively. At any time, the frequency of maximum intensity of component h is very nearly twice that of component f . In each of the examples marked, the two components start at frequencies of about 130 and 65 Mc/s respectively.

POLARIZATION MEASUREMENTS OF SOLAR RADIO BURSTS



POLARIZATION MEASUREMENTS OF SOLAR RADIO BURSTS



EVIDENCE OF ECHOES IN THE SOLAR CORONA FROM A NEW TYPE OF RADIO BURST

By J. A. ROBERTS*

[Manuscript received January 6, 1958]

Summary

A new spectral type of solar radio burst is described. The bursts contain two elements, the second being a repetition of the first after a delay of $1\frac{1}{2}$ –2 sec. In each element the frequency *increases* with time at a rate of 2–8 Mc/s per sec. The bursts are of very short duration and are confined to the longer metre wavelengths. Occasionally they occur within, and evidently form part of the structure of, a burst of spectral type III.

It is suggested that the second elements of the bursts are echoes of the first, reflected from lower levels of the solar corona. If the burst radiation is assumed to occur at the second harmonic of the coronal plasma frequency, the delay between the elements can be quantitatively explained providing the coronal density gradient is 1.5 times steeper than in the Baumbach-Allen model.

Two alternative explanations of the rising frequency characteristic are considered. Either the exciting disturbances travel in through the corona at speeds between 2 and 5×10^4 km sec⁻¹, or the outward travelling disturbances responsible for type III bursts encounter "hills" of electron density in the corona.

I. INTRODUCTION

It is known from the work of Wild and McCready (1950) that the dynamic frequency spectrum provides a natural means of classifying solar radio bursts. Furthermore, observations over the past 5 years with a spectrograph covering the range 40–240 Mc/s (Wild, Murray, and Rowe 1954) have shown that the great majority of solar bursts fall naturally into the three spectral classes defined by the former authors. Of the rarer events which do not fit this classification there are some which form a distinct class characterized by a double structure in which two short-lived features drift rapidly from lower to higher frequencies. This positive frequency drift is in contrast to the negative drift in spectral types II and III. These bursts are therefore termed *reverse drift pairs*, or more briefly, *reverse pairs*.

In the following section a detailed description of the reverse pairs is given under a number of headings, and in Section IV a possible theory of the origin of the bursts is considered. Section III contains a brief description of certain other bursts which have some features in common with the reverse pairs, but which, on present evidence, appear to form separate classes of phenomena.

A preliminary account of this work was presented by Dr. J. L. Pawsey at the U.R.S.I. Assembly in Boulder, Colorado, in August 1957.

* Division of Radiophysics, C.S.I.R.O., University Grounds, Chippendale, N.S.W.

II. A CLASS OF DOUBLE BURST WITH POSITIVE FREQUENCY DRIFT

The reverse drift pairs were first recognized in records taken on December 1 and 2, 1955. More than 20 of these quite characteristic bursts were observed at that time, and many hundreds of such bursts have been recorded since. There are therefore sufficient data available to delineate the basic characteristics of the type.

Spectral records showing examples of the bursts are reproduced in Plates 1 and 2, and in Figure 1 the appearance of the bursts on a single-frequency (40 Mc/s) record is compared with that of bursts of spectral types I and III.

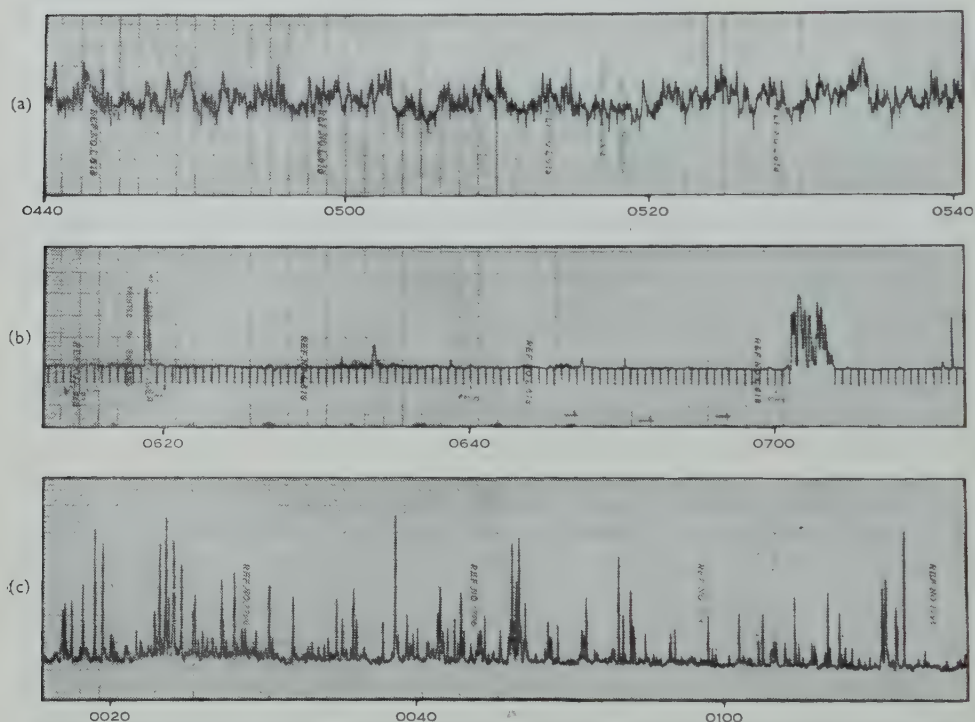


Fig. 1.—Single-frequency (40 Mc/s) records of (a) bursts of spectral type I (April 30, 1957), (b) bursts of spectral type III (April 4, 1957), (c) reverse drift pairs (March 13, 1957). In (b) time marks at intervals of $\frac{1}{2}$ min are shown by negative deflections (Universal times).

It is seen that in the frequency-time plane the reverse pairs have the appearance of two parallel ridges drifting rapidly from lower to higher frequencies at rates typically between 2 and 8 Mc/s per sec. The two elements of the bursts are usually very similar in form, although the intensities are sometimes markedly different. Many of the bursts begin at frequencies below 40 Mc/s, but from examples in which the full form of the burst is visible it is clear that both ridges commonly begin at the same frequency, the second ridge being delayed by approximately 2 sec after the first. In many cases the two ridges also terminate near the same frequency.

Each ridge is quite sharp—the duration at a single frequency is usually less than 1 sec, while the instantaneous bandwidth lies between 1 and 10 Mc/s. These durations and bandwidths are considerably less than those of type III bursts, which are the previously defined spectral type bearing most resemblance to the reverse pairs. A complete reverse pair may extend over a range of a few megacycles per second to a few tens of megacycles per second, with the corresponding total duration ranging from a few seconds to about 10 sec.

(a) *Occurrence*

Bursts of this type are rare. In the period from November 1955 to July 1957 observations were made on about 250 days and on only 38 of these days were reverse drift pairs observed. The bursts show a strong tendency to occur in storms lasting for hours or days. For the period of this investigation all clearly defined reverse pairs have been analysed—a total of 172 bursts in all. Of these, 59 (or 34 per cent.) occurred on March 12 and 13, 1957, a further 23 (13 per cent.) on June 5, 1957, 17 (10 per cent.) on December 1 and 2, 1955, and 10 (6 per cent.) on November 11, 1956. These four storms between them thus account for 63 per cent. of the bursts analysed.*

(b) *Frequency Range*

The frequency range of all the bursts studied is summarized in Figure 2. In this figure each burst is represented by a pair of contiguous lines which show the extent in frequency of the two elements of the burst. Figure 2(b) is a histogram showing the total number of bursts observed in each frequency interval.

It is immediately evident that the reverse pairs occur predominantly at the lower frequencies of the observed range, many extending below the frequency limit of the equipment at 40 Mc/s. The histogram shows an apparent decrease in the rate of occurrence for frequencies below 45 Mc/s, but this may be the result of overlooking inconspicuous events in which the reverse pairs barely extend into the frequency range of the observations.

(c) *Spacing between the Elements*

The double nature of the bursts and the constancy of the time separation of the two elements are distinctive characteristics of the reverse drift pairs. The histograms of Figure 3 show that for 80 per cent. of those analysed the time separation was from $1\frac{1}{2}$ to 2 sec. While reading errors certainly contribute to the spread of values in these figures, there are genuine variations between bursts, and between different frequencies in the same burst. Such variations with frequency are not always in one sense. The histograms show no significant consistent trend with frequency.

While the time separation of the elements lies between such close limits, the frequency separation varies considerably. This is evident from the wide variation in the rate of frequency drift (see Section II (f) below). Frequency separations range from a few megacycles per second to 10 Mc/s or more. The majority of values lie between 4 and 10 Mc/s.

* Since the preparation of this paper an outstanding storm of these bursts was recorded on September 2 and 3, 1957.

(d) *A Time Delay, not a Frequency Separation*

Double structure in the frequency-time plane could arise either from the simultaneous emission of two different bands of frequencies or from the occurrence of two similar events with a time delay between them. In the former case one

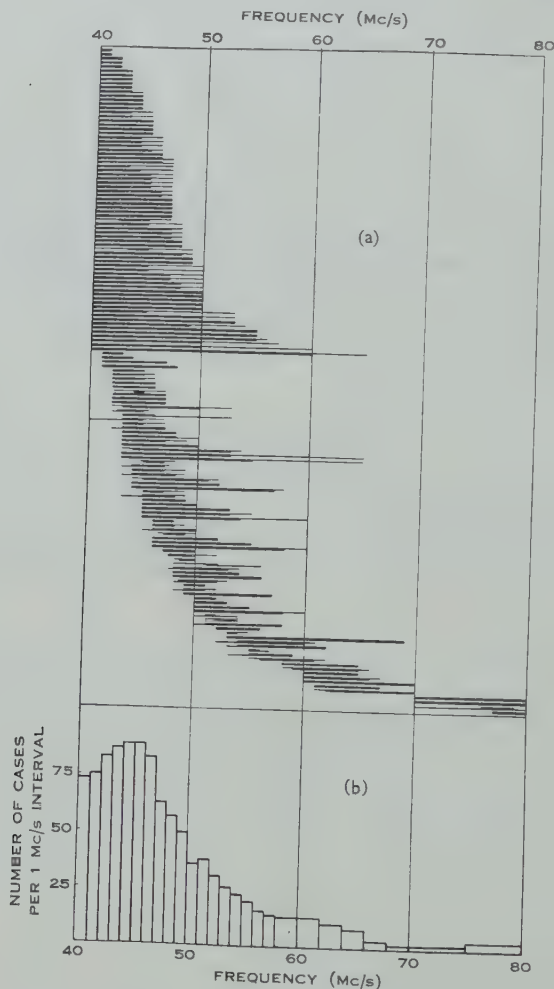


Fig. 2.—To illustrate the range of frequencies covered by the reverse pairs. (a) Each reverse pair is represented by a pair of contiguous lines which show the frequency extent of the two elements of the burst. (b) Histogram showing the prevalence of bursts at different frequencies.

would expect the burst to have the form shown in Figure 4(a), where the two elements begin (and end) simultaneously at different frequencies, and any structure in the burst occurs simultaneously in both elements. In the latter case the form shown in Figure 4(b) would be expected. Here both elements start at the same frequency (and both elements end at the same frequency), but

they are separated in time. Any structure in the first element is repeated in the second element after a time delay and *at the same frequency*.

Examination of the examples given in Plates 1 and 2 and of the diagrammatic representation in Figure 2 shows that not all the reverse pairs conform to either of the idealized sketches of Figure 4. However, a large number of the examples conform approximately to the model of Figure 4(b)—a separation in time. In

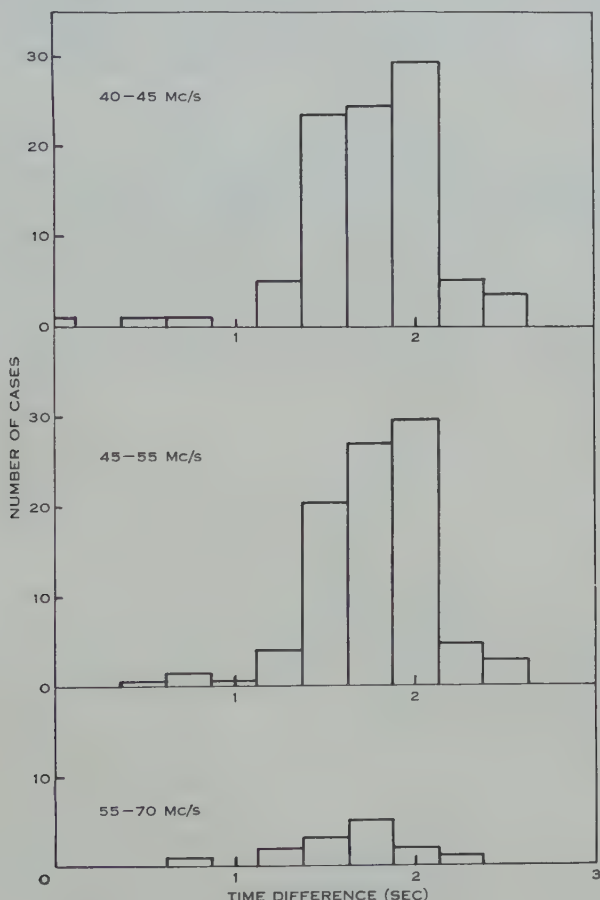


Fig. 3.—Histograms showing the separation in time of the two elements in the reverse pairs.

Plate 1 there are many cases of bursts in which both elements end at approximately the same frequency. Furthermore, in this plate, and particularly in Plate 2, there are striking examples of structure which is repeated in the second element of the burst at the same frequency (and not at the same time).

There is only one example in these figures which appears to conflict with the model of a time delay, namely the burst in Plate 1 on December 2, 1955 at 0032 U.T. In this reverse pair the curvature at the high frequencies appears to occur simultaneously in the two elements, rather than at the same frequency. As

no other clear example of this nature has been observed, it seems that this should be regarded as an unusual feature. A possible explanation is suggested in Section IV(e).

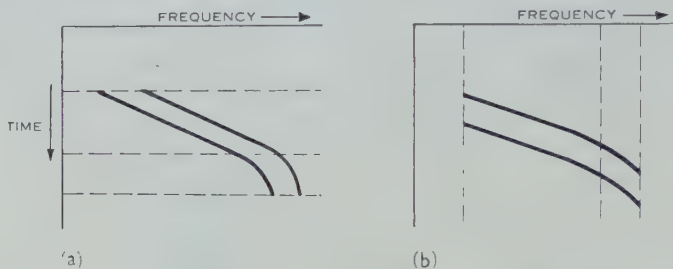


Fig. 4.—Idealized bursts. (a) Two bands of frequencies emitted simultaneously. (b) A single band of frequencies received via two channels with different propagation times.

The high degree to which the reverse pairs conform to the “time-delay” model is illustrated by Figure 5. From this it is seen that in the 83 bursts for which the beginning was observed, the difference between the frequencies at

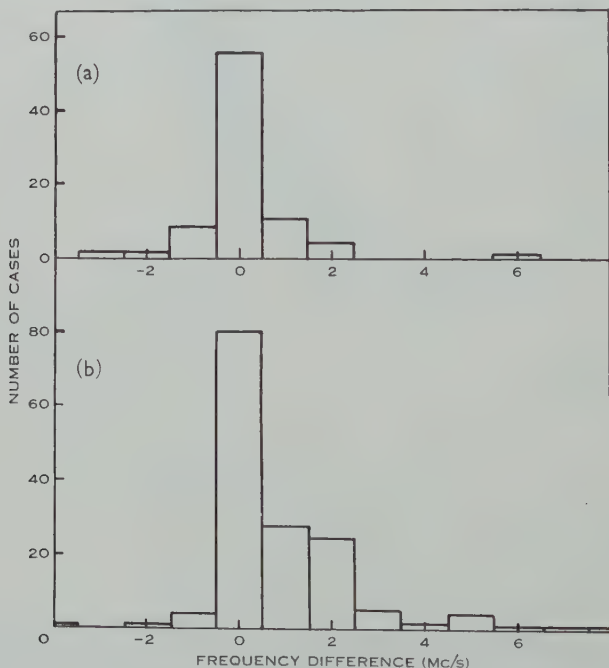


Fig. 5.—Histograms showing the difference in (a) the starting frequency, and (b) the finishing frequency of the two elements of the observed reverse pairs.

which the two parts began was distributed symmetrically about zero. Some 89 per cent. of the values lay between ± 1.5 Mc/s. If the bursts had in fact conformed to the double-frequency model (Fig. 4 (a)), these values would have

been negative and equal to the frequency separation of the bursts. Thus according to Section II(c) the values would have been mainly between -4 and -10 Mc/s.

Similar remarks apply to the distribution of the difference between the frequencies at which the two parts ended (Fig. 5 (b)). However, here there is a small, but definite, asymmetry in the distribution.

Finally, it may be mentioned that another feature favouring the time delay model is the small range of values of the time separation of the elements as compared with the frequency separation (Section II (c)).

(e) *Duration at a Single Frequency*

The reverse drift pairs appear on the records as very sharp features owing to their short duration at a single frequency. This duration is substantially independent of frequency. A histogram of durations given in Figure 6 shows that values greater than 1 sec are rare. By contrast, type III bursts at these frequencies usually last several seconds.

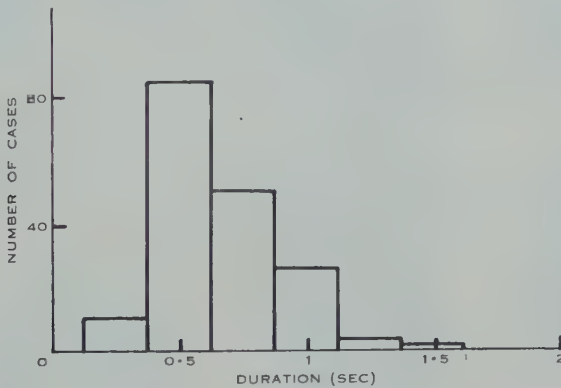


Fig. 6.—The duration of a single element of the reverse pairs at any one frequency.

Successive spectra are separated by $\frac{1}{2}$ sec in time, so that durations of less than $\frac{1}{2}$ sec are difficult to assess. For bursts with *total* durations of some seconds, however, the single-frequency duration may still be determined. It is likely that some of the values appearing in the histogram in the column centred on $\frac{1}{2}$ sec properly belong in the column centred on $\frac{1}{4}$ sec. Nevertheless, the commonest duration is probably close to $\frac{1}{2}$ sec.

(f) *Rate of Frequency Drift*

The distribution of the rate of change of frequency in the reverse pairs is shown in Figure 7. For 80 per cent. of the bursts the frequency increases at a rate between 2 and 8 Mc/s per sec. The frequency drift, besides being in the opposite sense to that in type III bursts, is also somewhat smaller in magnitude. For most type III bursts, at frequencies near 40 Mc/s the rate of decrease of the frequency of the fundamental band lies between 3 and 12 Mc/s per sec.

Not shown in Figure 7 are 13 cases in which the frequency drift was immeasurably great. In most of these cases the condition existed only at frequencies near 40 Mc/s and the rate decreased at higher frequencies. In two cases, however, all frequencies in the burst occurred simultaneously within the

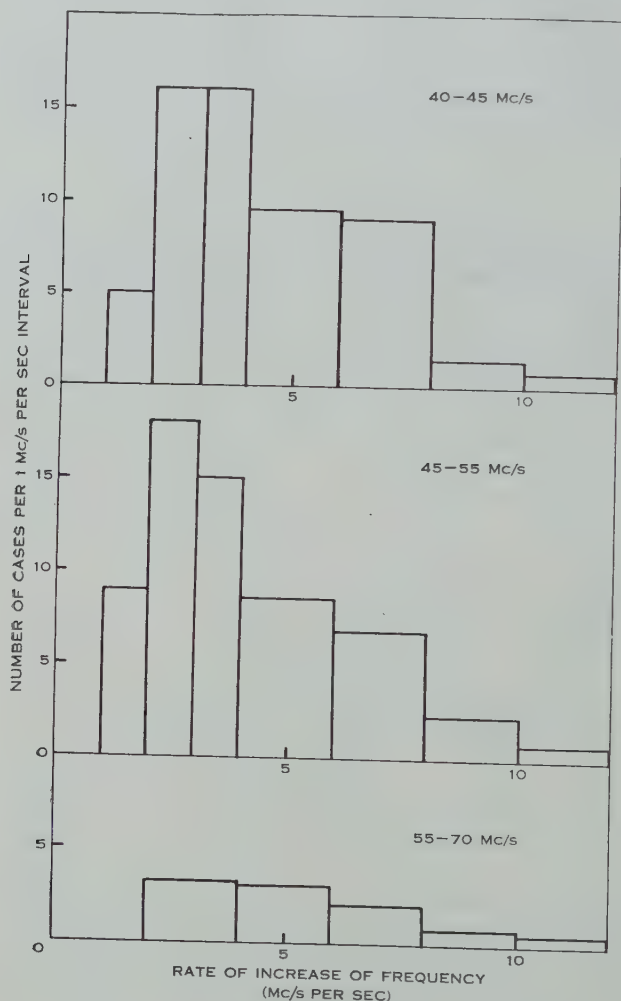
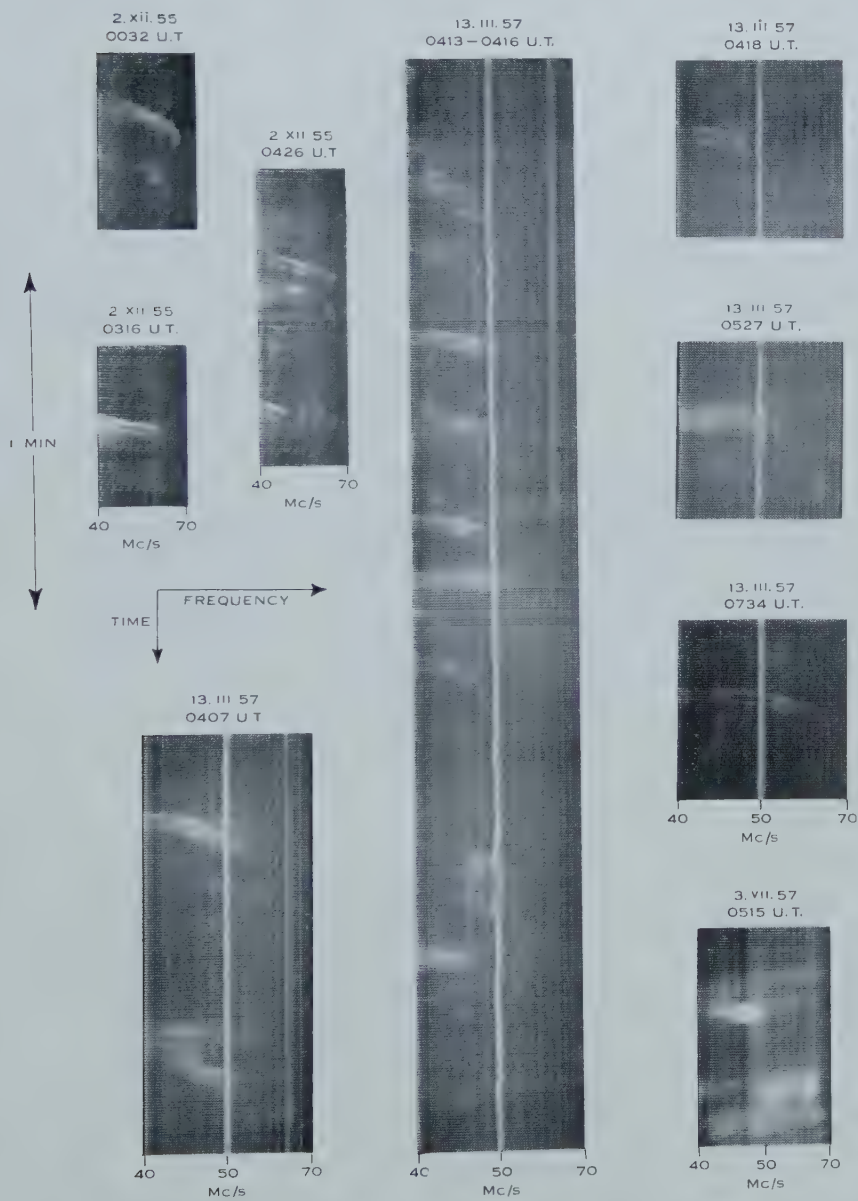


Fig. 7.—The rate of frequency drift in the reverse pairs.

limits of measurement. In addition, one case was noted in which the frequency drift was negative near 40 Mc/s but became positive at higher frequencies, so that frequencies near 42 Mc/s occurred first.

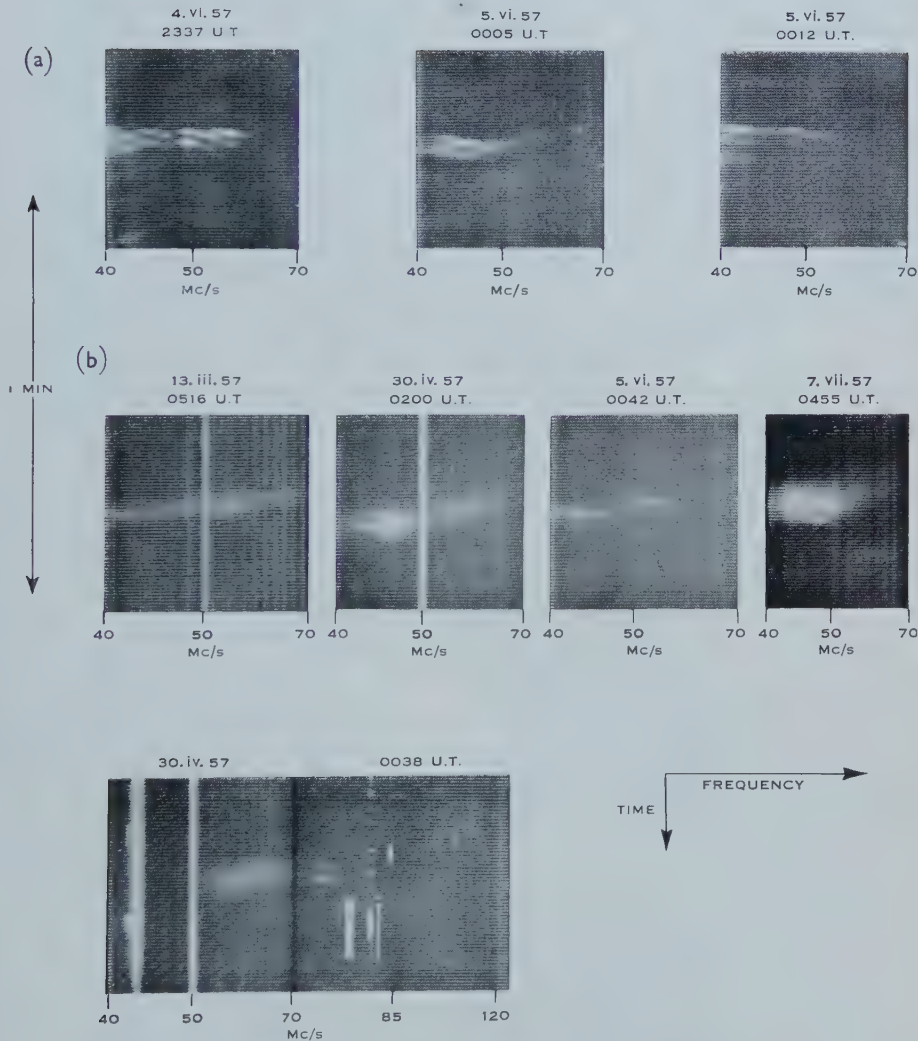
According to the histograms of Figure 7, the frequency drift, when averaged over many reverse pairs, is very similar at all frequencies. Most of the bursts in Plate 1 will be seen to agree individually with this statistical result, when allowance is made for the slight curvature produced by the compression of the frequency scale at the upper end. However, in some of the examples in this

A NEW TYPE OF SOLAR RADIO BURST



Reproductions of spectrum records showing double bursts with positive frequency drift (reverse drift pairs). Vertical features in these records are due to interfering signals or are instrumental and should be disregarded. The diagonal lines in the records of December 2, 1955 are produced by interference modulated at 100 c/s.

A NEW TYPE OF SOLAR RADIO BURST



Further examples of reverse drift pairs. (a) Examples showing sudden changes in the rate of frequency drift. (b) Reverse pairs occurring within bursts of spectral Type III.

A NEW TYPE OF SOLAR RADIO BURST

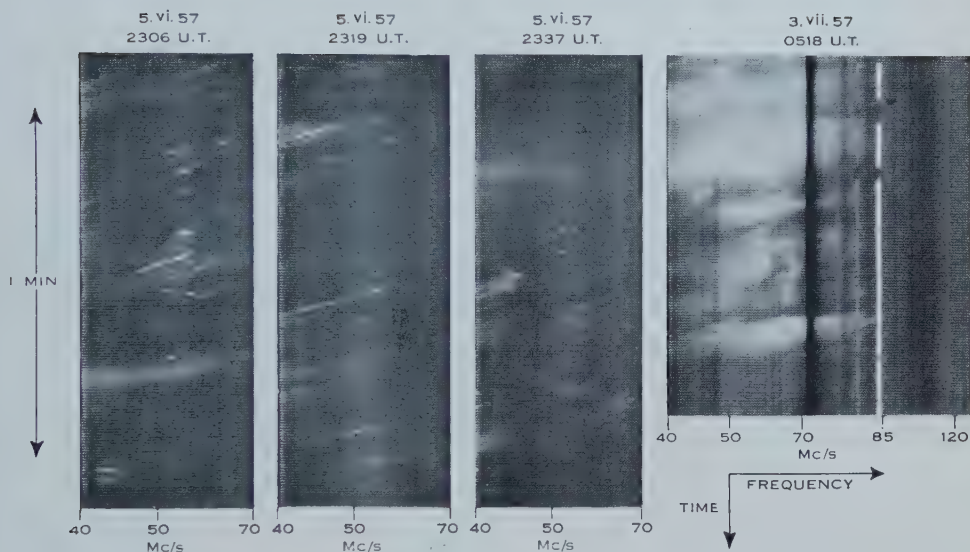


Fig. 1.—Reproductions of spectrum records of bursts of very short duration with negative frequency drift. Vertical features in these records are due to interfering signals or are instrumental and should be disregarded.

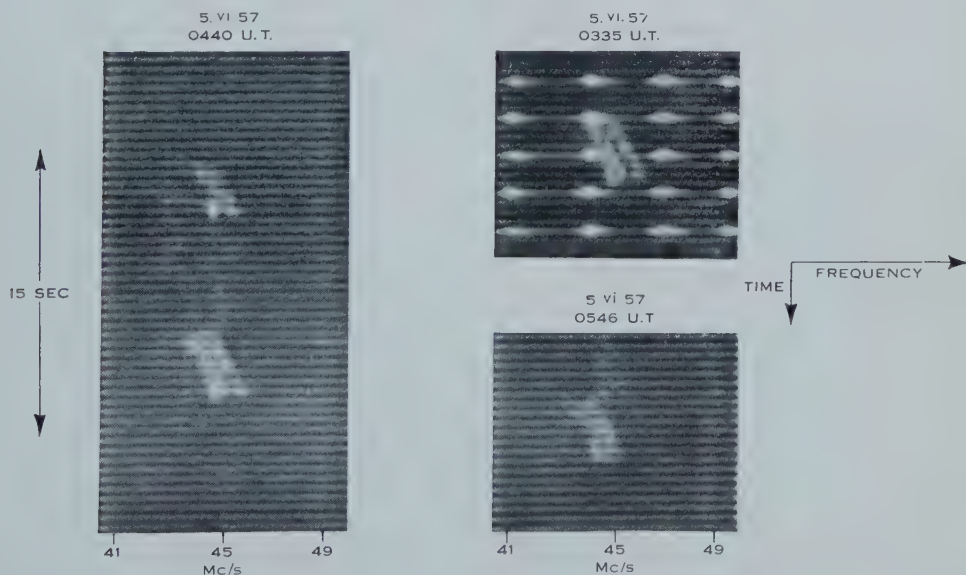


Fig. 2.—Spectrum records of a unique type of double burst with positive frequency drift recorded on June 5, 1957. In the record at 0335 U.T. frequency markers are visible at harmonics of 2.74 Mc/s.

plate the rate of change of frequency decreases markedly at the higher frequencies (later times). In the examples in Plate 2 (a) there are sudden changes in the rate of frequency drift, several changes occurring in the same burst.

(g) *Intensity and Polarization*

Compared with other bursts occurring at these frequencies, the reverse pairs are relatively weak. Measured flux densities in one plane of polarization range from the limit of detection ($\sim 5 \times 10^{-21} \text{W m}^{-2} (\text{c/s})^{-1}$) up to about $5 \times 10^{-20} \text{W m}^{-2} (\text{c/s})^{-1}$. Often the two elements of the bursts are of similar intensity, but there is a tendency for the second element to be weaker. This is illustrated by the histogram of intensity ratios in Figure 8, which summarizes the data for 26

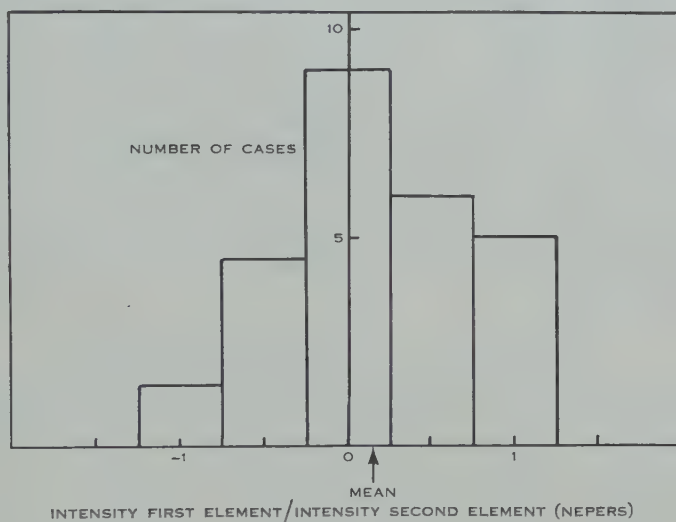


Fig. 8.—Histogram showing the ratio of the intensities of the two elements of the reverse pairs. Data for the 26 strongest bursts in three periods of approximately 15 min each on March 12 and 13, 1957. Results derived by microphotometering the film records, scanning along a line of constant frequency, which was 40 Mc/s for two of the periods and 45 Mc/s for the third period.

reverse pairs on March 12 and 13, 1957. This histogram is thought to be reasonably representative of this storm. On other occasions the relative weakness of the second elements has been more marked; numerous cases of single elements bursts have been noticed.

As yet no observations have been made of the polarization of these bursts.*

(h) *Positional Observations*

On March 12 and 13, 1957 some reverse pairs were observed with a swept-frequency interferometer similar to that described by Wild and Roberts (1956). When the approximate declination of a source is known, this instrument is capable of determining the Right Ascension of the source to within a few minutes

* Reverse drift pairs observed on September 19, 1957, were found to be not strongly polarized.

of arc. Unfortunately the reverse pairs observed with the interferometer did not extend over a sufficiently wide range of frequencies to enable all ambiguities to be resolved. There are therefore a number of possible values of Right Ascension for the bursts. If the declination of the source is assumed to be close to that of the Sun, all these position in Right Ascension lie within a solar diameter of the centre of the disk, and the most probable position is 5 min of arc E. These observations therefore strongly suggest that the reverse pairs are indeed of solar origin, and this conclusion is completely supported by the association with type III bursts, which is discussed in the next subsection.

On June 4 and 5, 1956, Kraus (1956a) observed short duration "double-humped" bursts at 27 Mc/s. He gave an indirect argument suggesting that the bursts originated on Venus. As the spectrograph recorded a small storm of reverse drift pairs on June 5, it seems possible that the bursts reported by Kraus may be of the same nature.* The interferometer records referred to above were therefore re-examined to see if it were possible that the reverse pairs came from Venus. Although, on the day in question, Venus was within 10° of the Sun, the interferometer records make it quite impossible that these bursts could have come from Venus. For a source on Venus the frequency separation of the interferometer lobes would have been 32 Mc/s while the observed value was 2.73 Mc/s. Similar considerations apply in the case of Mercury, the only other planet close to the Sun at this period.

(i) *Relationship to Bursts of Other Spectral Types*

About 10 per cent. of the reverse pairs appear to occur in, and evidently form part of, a burst of spectral type III. Some examples of such events are shown in Plate 2(b), and the last example in Plate 1 may also be of this type. These reverse pairs occurring within type III bursts appear to be otherwise indistinguishable from those occurring separately. However, as yet no reverse pair extending over a wide range of frequencies (> 15 Mc/s) has been observed in a type III burst.

In all of these events the reverse pair is of greater intensity than the type III burst, and in most cases the type III event is very weak and diffuse. A reverse pair has never been observed in a very intense type III burst.

The number of clear examples of these events so far recorded is insufficient to define properly the relationship between the two bursts. As is seen from Plate 2(b) there seems to be a tendency for the extent of the first element of the reverse pair to be limited to the extent of the type III burst. However, in other less clearly defined cases, the reverse pair appears to extend beyond the type III burst. In some cases the rate of frequency drift in the type III burst is apparently smaller on the low frequency side of the reverse pair.

No definite relationship has been found with any other type of activity, and indeed on a number of occasions (e.g. March 12, 1957) reverse pairs have been

* It is not yet possible to say whether the reverse drift pairs are modulated at audio frequencies as described by Kraus. In a later publication Kraus (1956b) found that double-humped bursts on June 29, 1956 had a *negative* frequency drift. No spectrograph records are available for this date, but see Section III (a) of the present paper.

the only activity recorded. However, in the case of all the four storms mentioned in Section II (a), a type I storm was also in progress at the higher frequencies for at least part of the time.

(j) *Repetition of Features at Several Frequencies*

Another feature of the reverse pairs which may be of some significance is the occasional occurrence of an event in which two or more reverse pairs occur simultaneously, or nearly simultaneously, at different frequencies. Some examples are seen in Plate 2. The ratio of the frequencies of the two bursts is considerably less than 2 and in fact varies widely from case to case. In the examples in Plate 2(b) the two related reverse pairs are evidently superimposed on the same type III burst. Other examples of this behaviour are known. In Plate 2(a) there is a quite extraordinary example of this repetition. The burst appears to consist of four separate reverse pairs, together with a superimposed pair with negative frequency drift which extends from below 40 Mc/s to 44 Mc/s. These bursts are so combined that over the greater part of the frequency range the event appears to consist of two ridges parallel to the frequency axis.

III. OTHER MULTIPLE BURSTS

In this section a brief description is given of two other types of multiple burst which are similar in some respects to the reverse drift pairs described above.

(a) *Bursts with Negative Frequency Drift*

A small number of bursts have now been recorded which are very similar to the reverse pairs except that the sense of frequency drift is opposite, i.e. the frequency decreases with the passage of time. Some examples of these bursts may be seen in Plate 3, Figure 1. These bursts are even more rare than the reverse pairs. Only 12 bursts have been examined in detail, so that the description given here is necessarily tentative.

The bursts all occurred on days when reverse pairs were also recorded. The spacing of the elements again appeared to be mainly a separation in time; the mean separation measured was $1\frac{1}{2}$ sec and the range of values $1\frac{1}{4}$ –2 sec. In the sample available the duration of each element of the burst (at a single frequency) was noticeably less than in the reverse pairs. In most of the examples in Figure 1 of Plate 3 the durations are less than $\frac{1}{2}$ sec so that at any one frequency the bursts appear on only one scan.

Frequency drift rates range from -1.5 to -3 Mc/s per sec, with a mean value of -2.4 Mc/s per sec. These values are very similar in *magnitude* to those for the reverse pairs (Fig. 7): the rates are noticeably lower than those for type III bursts. The contrast between these bursts and type III bursts is well illustrated by the last example in Figure 1 of Plate 3 where two of these double bursts with negative frequency drift appear superimposed on a type III burst.

All the bursts of this nature so far recorded have occurred at frequencies below 80 Mc/s, but the concentration to the lower frequencies does not appear to be as marked as in the case of the reverse pairs. None of the bursts recorded

has been of very great intensity. If the sample at present available is typical, these bursts do not form such a homogeneous class as the reverse pairs. The sample contains a high proportion of single bursts and some triple bursts.

(b) *Unusual Bursts with Positive Drift*

On June 5, 1957, five unusual double bursts were recorded. All were remarkably similar in form. Four are reproduced in Figure 2 of Plate 3. These bursts may be described as double bursts with positive frequency drift, but they do not appear to belong to the class referred to as reverse drift pairs. In the bursts of Plate 3, Figure 2, the two elements begin at the same time and end at the same time, i.e. the splitting is in frequency and not in time. Furthermore, the rate of frequency drift is very much less than in the reverse pairs—the values range from 0.33 to 0.42 Mc/s per sec.

All five bursts occurred between 43 and 46 Mc/s and each lasted between 3 and 4 sec. The frequency separation of the two elements ranged from $\frac{1}{2}$ to 1 Mc/s. The two parts of the bursts are not simple ridges, but contain remarkable fine structure, some of which is visible in the figure. In several of the bursts the lower frequency element is double, with a separation of a few tenths of a megacycle per second between the “fine-structure” ridges. This separation varies between bursts and throughout the course of any one burst. In some cases the high-frequency element also appears to be split, the separation in this case being very much less (~ 0.03 Mc/s).

It is necessary to add that there is no proof that these bursts are of solar origin. Many reverse drift pairs were also observed on this day, but, as the positions of these unusual bursts were not measured, it is possible that they did not come from the Sun.

IV. A THEORY INVOLVING ECHOES IN THE REGULAR CORONA

(a) *Introduction*

This section outlines a possible theory of the origin of the double bursts with positive frequency drift, described in Section II. The discussion is confined to these reverse drift pairs as the present observations appear to be inadequate to define the basic characteristics of the other bursts described in Section III.

(i) *Echoes in the Regular Corona.*—It was shown in Section II(d) that the second elements of the reverse pairs are approximate repetitions of the first elements after a time delay of $1\frac{1}{2}$ –2 sec. This suggests that a single band of frequencies is generated but that it is received via two paths with different propagation times. Many possibilities of this nature suggest themselves—echoes could occur from coronal irregularities, propagation could be via the two magneto-ionic modes, etc. However, echoes from the *regular* corona would seem to be the first possibility to investigate, since it has been shown (Jaeger and Westfold 1950) that such echoes should occur and should have time delays of several seconds.*

* Jaeger and Westfold in fact proposed that the “double-humped” forms of “unpolarized bursts” were due to echoes from the regular corona. However, it is now known that in these bursts (of spectral type III) the second part is the second harmonic and not a delayed repetition of the first part (Wild, Murray, and Rowe 1954).

Jaeger and Westfold showed that there are, in general, two rays to the Earth from any accessible point in the solar corona. These may be described as the direct and reflected rays (Fig. 9). The time delay between receiving a disturbance via the direct and reflected paths—the echo delay—depends on the height of the source, its position on the disk, and the frequency. When the source lies on the escape level for the frequency in question the two paths are identical and the delay is zero. As the height of the source is increased above the escape level the echo delay increases, reaching 5–10 sec at a height of 1 solar radius. Thus, to explain the observed delays in terms of echoes, we seek a natural reason for locating the source at such a height as to produce delays of the order of 2 sec.

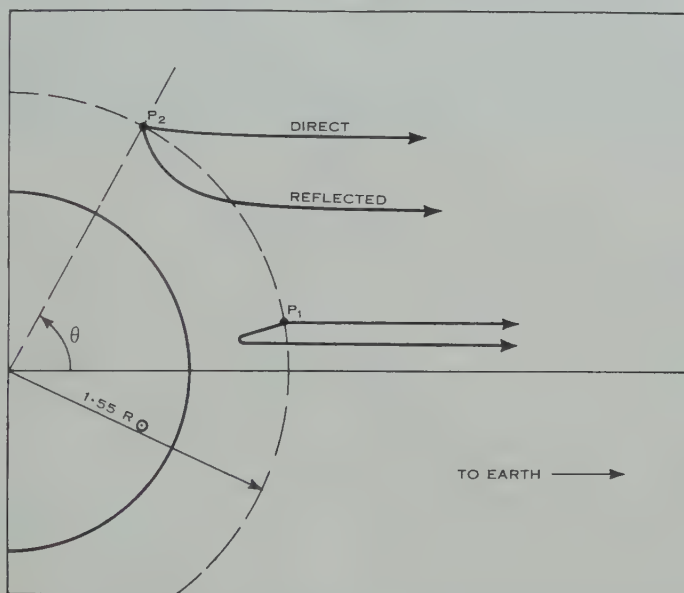


Fig. 9.—Showing the paths of direct and reflected rays in the corona for a source located at the 30 Mc/s plasma level and radiating at 60 Mc/s.

(ii) *Radiation at the Second Harmonic of the Coronal Plasma Frequency.*—The theory considered here supposes that the height of origin is determined by the condition that the radiation occurs at frequencies close to the second harmonic of the local coronal plasma frequency. This hypothesis is a simple extension of the theory which has found support as an explanation of the features of bursts of spectral types II and III (Wild, Murray, and Rowe 1954). It accounts naturally for the constancy of the time delay over periods of years and it will be shown that it can give reasonable quantitative agreement with the observations.

For a source near the centre of the disk radiation at frequencies near the *fundamental* plasma frequency would also be expected to escape (Wild, Murray, and Rowe 1954). For these frequencies the echo delay would be only a fraction of a second, so that such a burst would appear single. This leads to one of the tests of the present hypothesis discussed in Section IV (c) (iv).

(iii) *The Positive Frequency Drift.*—Two explanations of the positive frequency drift are considered. In the first it is supposed that a disturbance travels inward through the corona exciting successively higher frequencies of oscillation as it penetrates to regions of higher electron density. On this theory observed frequency drift rates imply speeds of $2\text{--}5 \times 10^4$ km sec⁻¹, which are far greater than the speed resulting from falling under gravity from infinity.

In the other theory, suggested to the author by Mr. J. P. Wild of this laboratory, the radiation is supposed to be produced by outward travelling disturbances when they encounter "hills" of electron density in the corona. This second explanation is suggested by the occurrence of these bursts within type III bursts (Section II (i)), the type III bursts being attributed to the interaction of the outward moving disturbances with the regular corona. The common occurrence of the reverse pairs without an accompanying type III burst requires one to assume that in these cases the type III part is too weak to be detected. On this theory the frequency drift rates imply that the gradients of electron density in the "hills" are about two-thirds of the gradient in the regular corona.

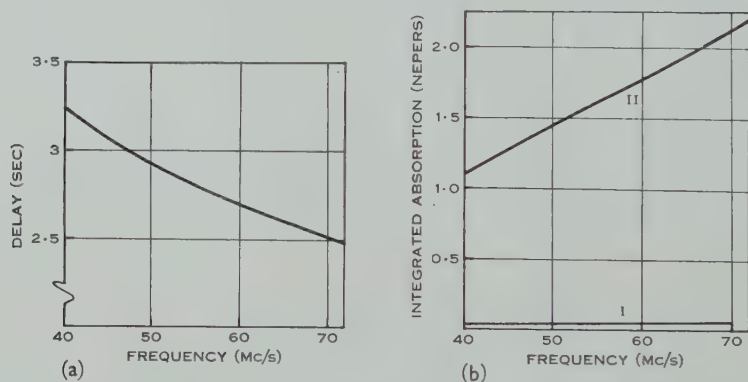


Fig. 10.—Frequency dependence of the properties of rays from a source in the centre of the disk radiating at twice the local coronal plasma frequency. Baumbach-Allen model. (a) Time delay of the reflected ray after the direct ray. (b) Integrated absorption in the direct ray (curve I) and excess integrated absorption in the reflected ray (curve II).

(b) Predicted Properties of the Echoes

Many of the characteristics of the echoes from a source in the corona radiating at the second harmonic of the local plasma frequency can be deduced from the curves given by Jaeger and Westfold (1950). Some further properties are derived in Appendix I, and the results are summarized in Figures 10 and 11. In the derivation of these results it has been assumed that the corona is spherically symmetrical and that the electron density distribution is that given by Allen (1947).

Figure 10 shows the frequency dependence of the delay between the direct and reflected waves for a source in the centre of the disk. The absorption in the direct ray and the excess absorption in the reflected ray are also given in

this figure. If the source is not in the centre of the disk the predicted delay is less (Fig. 11 (a)). As the source is displaced from the centre of the disk (but kept at the same height) the paths of the two rays become more alike, and when the source lies on the escape level the paths are identical and the time delay is zero. No radiation can be received from sources beyond this point.

The absorption in the direct ray increases as the source is displaced from the centre of the disk, but is never appreciable (Fig. 11 (b)). The excess absorption in the reflected ray decreases to zero as the two paths approach identity.

Refraction modifies the emission polar diagram of the source (Fig. 11 (c)). When the source is near the centre of the disk the atmosphere acts as a converging lens for the direct rays but the reflected rays diverge so that the reflected image is of lower apparent intensity. As the source is displaced from the centre of the disk both direct and reflected rays become converging, so that if rays leave the source isotropically the apparent intensity of the source increases. As the source approaches the escape level this focusing effect becomes very pronounced.

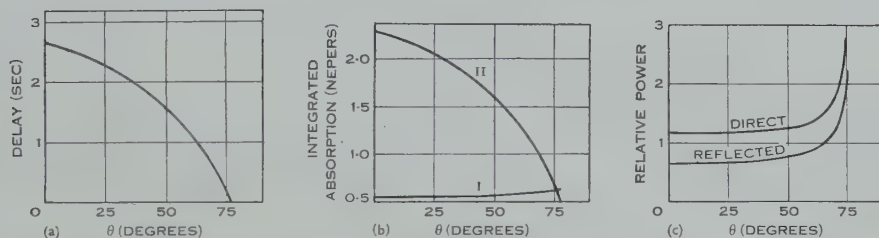


Fig. 11.—Centre to limb variation for a source at the 30 Mc/s plasma level radiating at a frequency of 60 Mc/s. The source position angle θ is defined in Figure 9. (a) Time delay of the reflected ray after the direct ray. (b) Integrated absorption in the direct ray (curve I) and excess integrated absorption in the reflected ray (curve II). (c) Effective emission polar diagram (unnormalized). The results in this figure are derived from the curves given by Jaeger and Westfold (1950). The effective emission polar diagram in (c) is just $d\theta'/d\theta$ where θ' , θ are the initial and final directions of the ray measured from a radial line through the source.

(c) Comparison with Observations

(i) *Time Delays.*—A comparison of the observed time delays (Fig. 3) with those predicted for central rays (Fig. 10(a)) shows that the predicted values are certainly of the right order. However, if the sources were observed at all (accessible) distances from the centre of the disk, the predicted delays would range over all values from zero up to that for central rays (Fig. 11(a)). This conflicts with the narrow range of values observed. If this theory is correct, therefore, it seems that the sources are observed only over a narrow range of distances from the disk centre. Neither absorption (Fig. 11 (b)) nor refraction (Fig. 11 (c)) can account for this, so that it is necessary to suppose that the rays leave the source preferentially in certain directions.

If the sources are observed only near the centre of the disk, i.e. if the preferential directions are radial (inwards and outwards) and the sources lie at low latitudes, then the predicted delays are too great by approximately 1 sec (Fig. 10(a)). To secure agreement it is then necessary to suppose that the density

gradient in the corona is greater, by a factor of about 1.5, than that given by the Baumbach-Allen model (r^{-6} approximation).

Alternatively the observed delays would agree with those predicted for the Baumbach-Allen model if the sources were observed only between about 40 and 60° from the centre of the disk. Such a restriction to a non-central range of angles would arise if the sources occurred at high latitudes ($\sim 40^\circ$), and had radially directed polar diagrams of appropriate width.

It is of course necessary to consider whether the deficit of low values in the measured delays (Fig. 3) is not due to the coalescence of the two elements when the delay becomes small. Although this effect should become appreciable for time separations ~ 1 sec, it is not thought that many bursts are overlooked through this process as relatively few single bursts of long duration are observed.

(ii) *Differential Absorption*.—The relative intensities of the two elements of the bursts should provide another parameter for testing the theory. However, there are insufficient measurements available for an exhaustive test. Furthermore, the measurements on March 12 and 13, 1957 (Fig. 8) show a very wide scatter, including cases in which the second element is the stronger. This suggests that some other process, perhaps scattering by irregularities, produces relative intensity differences between the two rays, which can be of either sign.

Figure 8 shows that the *mean* intensity ratio on March 12 and 13, 1957 was only about one-sixth of the predicted ratio for central rays (Fig. 10(b)).

(iii) *The Two Explanations of the Positive Frequency Drift*.—It is not possible at present to decide between the two suggested explanations of the positive frequency drift, namely, inward-travelling disturbances exciting the regular corona or outward-travelling disturbances encountering “hills” of electron density in the corona. However, the former explanation is slightly favoured, as the short duration and narrow bandwidth are then readily explained merely by supposing that the disturbances are of sufficiently small extent ($\lesssim 10^4$ km). With the alternative theory it is difficult to account for the narrow bandwidth. The appearance of the reverse pairs superimposed on the type III bursts, without evident distortion of the type III event, implies that the lateral extent of the travelling disturbance is very much greater than that of the coronal “hill”. One would therefore expect the disturbance to excite a continuous range of plasma frequencies corresponding to the range of densities in the hill up to the highest point reached by the disturbance.

Further study of the relationship between the reverse pairs and type III bursts might well decide between the two theories. While the occurrence of reverse pairs within type III bursts is inherent in the alternative theory, with the first theory such an event must imply a “collision” of an inward and an outward moving disturbance. There are thus differences between the predictions of the two theories regarding the details of such an event. While the form of the event observed in a number of cases (Plate 2(b)) favours the collision model, there are considerable variations between events. Many more examples will be needed to define the properties of this association adequately.

(iv) *The Second Harmonic Nature of the Bursts.*—As was pointed out in Section IV(a) (ii) the theory discussed here implies that reverse pairs originating near the centre of the disk will be accompanied by *single* bursts—the fundamental band—at half the frequency. This prediction should provide a test of the theory, provided the bursts sometimes originate near the centre of the disk (cf. Section IV (c) (i)). So far only one double burst has been observed at a sufficiently high frequency for the predicted fundamental band to lie within the frequency range of the equipment. In this case no fundamental was detected. However, this burst was by no means typical of the class; the time separation was $4\frac{1}{2}$ sec and the duration at a single frequency was 2 sec.

This test may be made in another way if a plasma origin is assumed for the type III bursts also. For then it follows that the type III band in which the double bursts occur should be the second harmonic band. Unfortunately, in no event of this kind so far recorded has it been possible to assign a harmonic number to the type III component involved.

It appears that further observations, preferably extending to lower frequencies, are needed to check these predictions.

(d) *Echoes of Other Solar Bursts*

It remains to consider whether echoes of other solar bursts should not be observed if this theory were correct. Evidently the second harmonic components of type III bursts are all that need be considered. At frequencies near 40 Mc/s the durations of these bursts are many seconds, so that any echo would be masked. At higher frequencies, the situation is more confused. Above about 100 Mc/s the durations are often 1 sec or less, so that on the model of Figure 10 separate echoes should be recorded. A number of examples of possible echoes are found in the records, but the tendency for type III bursts to occur together in groups makes it impossible to be sure that the suspected echoes are not, in fact, separate bursts.

(e) *Structure Simultaneous in Time*

In Section II(d) attention was drawn to an exceptional case in Plate 1 in which the curvature in the two elements appears to occur more nearly at the same time than at the same frequency. As there seems to be ample evidence to show that the second element of a reverse pair is usually a delayed repetition of the first, it is suggested that in this case both elements were simultaneously subjected to some extra delay. Such a delay could perhaps arise from the source (and the image) passing behind a coronal streamer.

(f) *The Decay Time of Coronal Plasma Oscillations*

Attention may be drawn to a general conclusion which follows from the assumption that the reverse pairs and type III bursts both originate in coronal plasma oscillations. Westfold (1949) (see also Payne-Scott 1949) suggested that the duration of type III bursts (at one frequency) was determined by the decay of the oscillations of the coronal plasma after the excitation was removed. As the duration of the present bursts is very much less than that of type III bursts (at the same frequency), it follows that the duration of type III bursts is

not determined by the decay time; it is presumably determined by the extent of the exciting disturbance.

In proposing the decay-time theory of the duration of type III bursts, Westfold showed that the durations were close to the values predicted on the assumption that the oscillations were damped by collisions. However, it has been suggested by Dr. R. Q. Twiss (personal communication) that coronal plasma oscillations would be much more rapidly damped by the diffusing effect of thermal velocities—a process known as Landau damping (Landau 1946; Berz 1956).

V. CONCLUSION

The main characteristics which distinguish the reverse drift pairs from other known types of solar radio bursts may be summarized as follows:

- (i) The bursts are double, the second part apparently being a repetition of the first part after a delay of $1\frac{1}{2}$ –2 sec (Section II(c) and (d)).
- (ii) The frequency drift is positive, the frequency increasing with time at a rate of 2–8 Mc/s per sec (Section II(f)).
- (iii) The duration of the bursts at a single frequency is very short, usually less than 1 sec (Section II(e)).
- (iv) The bursts are confined to the lower frequencies, the rate of occurrence of the bursts increasing steadily with decreasing frequency, at least to ~ 40 Mc/s (Section II(b)).
- (v) Occasionally the bursts appear in close association with type III bursts and may be a feature of those bursts.

All of these characteristics distinguish the bursts as a separate type of activity—they are not, for example, a variety of type III burst with positive frequency drift. However, there is a type of burst with negative frequency drift, which resembles the reverse pairs fairly closely (Section III(a)).

The theory discussed in Section IV appears to be capable of explaining the main features of these bursts, but to achieve quantitative agreement some additional assumption or modification is required (Section IV(c) (i) and (ii)). With certain assumptions as to the emission polar diagram of the source, reasonable agreement could be secured if the gradient of electron density in the corona were steeper (by a factor ~ 1.5) than that given by Allen (1947). Alternatively agreement may be possible if the bursts occurred at fairly high latitudes ($\sim 40^\circ$).

This theory predicts that the reverse pairs will sometimes be accompanied by *single* bursts at half their frequency, and that they will not occur in the fundamental band of type III bursts. Further observations, particularly if extended to lower frequencies, should allow these predictions to be tested. Such observations will also define more clearly the relationship of the reverse pairs to the type III bursts in which they occur. It is felt that a study of this relationship will throw considerable light on the origin of the bursts. Accurate position measurements would provide a further check of the theory, and, if both coordinates of the burst position were measured, a decision between some of the alternative forms of the theory would be possible.

Whether or not the suggested theory is correct, it can be concluded from propagation conditions alone that the reverse pairs are generated in the high corona. This itself is suggestive of incoming material. No optical counterpart to the bursts has been discovered; these considerations suggest that correlation should be sought with optical phenomena in the upper corona, and perhaps at high latitudes.

VI. ACKNOWLEDGMENTS

An investigation of this nature, in which rare events are studied, requires the cooperation of many persons to secure a long series of records. The investigation forms part of the C.S.I.R.O. Radiophysics Laboratory's project for the study of the spectra of solar radio disturbances. This project is under the direction of Mr. J. P. Wild, who, with Mr. J. D. Murray, was responsible for the design of the spectrograph. Mr. K. V. Sheridan, Mr. M. Komesaroff, and Mr. J. Joissee have contributed largely to the maintenance and running of the equipment, and Mrs. P. Hodgson has assisted with the analysis of the records.

The author has to thank Mr. J. P. Wild, Mr. S. F. Smerd, and Dr. R. Q. Twiss for many stimulating discussions of possible theories of the generation of these bursts.

VII. REFERENCES

- ALLEN, C. W. (1947).—*Mon. Not. R. Astr. Soc.* **107**: 426.
 BERZ, F. (1956).—*Proc. Phys. Soc. Lond.* B **69**: 939.
 JAEGER, J. C., and WESTFOLD, K. C. (1950).—*Aust. J. Sci. Res.* A **3**: 376.
 KRAUS, J. D. (1956a).—*Nature* **178**: 159.
 KRAUS, J. D. (1956b).—Ohio State Univ. Radio Astronomy Rep. No. 6.
 LANDAU, L. (1946).—*J. Phys. U.S.S.R.* **10**: 25.
 PAYNE-SCOTT, RUBY (1949).—*Aust. J. Sci. Res.* A **2**: 214.
 WESTFOLD, K. C. (1949).—*Aust. J. Sci. Res.* A **2**: 169.
 WILD, J. P., and MCCREADY, L. L. (1950).—*Aust. J. Sci. Res.* A **3**: 387.
 WILD, J. P., MURRAY, J. D., and ROWE, W. C. (1954).—*Aust. J. Phys.* **7**: 439.
 WILD, J. P., and ROBERTS, J. A. (1956).—*J. Atmos. Terr. Phys.* **8**: 55.

APPENDIX I

The frequency dependence of the echo delay for central rays is derived under the following assumptions:

(i) The radiation occurs at the second harmonic of the local coronal plasma frequency,

(ii) The coronal electron density is given by the r^{-6} approximation to the Baumbach-Allen distribution, and

(iii) Coronal magnetic fields do not produce appreciable effects.

The refractive index μ is then given by (Jaeger and Westfold 1950, equation (10))

$$\mu^2 = 1 - 1.41 \times 10^9 f^{-2} r^{-6},$$

where f is the wave frequency in megacycles per second and r is the distance from the centre of the Sun in units of 10^5 km. It will be convenient to introduce $x = 1.41 \times 10^9 f^{-2} r^{-6}$, so that $\mu^2 = 1 - x$.

When the wave frequency is equal to the local plasma frequency $x=1$ and $\mu=0$. This is the turning point for a central ray. At the generation level, where the frequency is twice the local plasma frequency, $x=\frac{1}{4}$. The time delay between the direct and reflected rays is thus given by

$$2 \int_{x=1}^{x=\frac{1}{4}} \frac{dr}{\mu c},$$

where c is the free-space velocity of light.

The integral is evaluated by first changing the variable to give

$$3 \cdot 72 f^{-1/3} \int_{\frac{1}{4}}^1 x^{-7/6} (1-x)^{-1/6} dx.$$

Then expanding $(1-x)^{-1/6}$ and integrating term by term gives

$$3 \cdot 72 f^{-1/3} \left[\sum_{n=0}^{\infty} \frac{(2n)! x^{n-1/6}}{n! n! 2^{2n} (n-1/6)} \right]_{\frac{1}{4}}^1.$$

The series converges for $x \leq 1$. At the upper limit ($x=1$) the sum is known to be

$$\frac{\Gamma(-1/6)\Gamma(\frac{1}{2})}{\Gamma(\frac{1}{2}-1/6)} = -4 \cdot 23.$$

At the lower limit the convergence is reasonably rapid and the sum is found to be approximately $-7 \cdot 35$.

Thus one finds for the time delay between the direct and reflected rays

$$11 \cdot 6 f^{-1/3} \text{ sec},$$

where f is the wave frequency in Mc/s. These values are plotted in Figure 10 (a).

The integrated absorption for central rays is computed in a similar fashion. The integrated absorption is

$$2 \int \kappa dr,$$

where κ is the *field* attenuation coefficient. Using Jaeger and Westfold's value for the collision frequency (but dropping the r^{-10} terms), the integral can again be expressed in terms of x to give

$$2 \int \kappa dr = 4 \cdot 61 \times 10^{-3} f^{7/6} \int x^{7/12} (1-x)^{-1/6} dx.$$

The integral is evaluated by expanding $(1-x)^{-1/6}$ and one finds ultimately for the integrated absorption in the direct ray

$$3 \cdot 5 \times 10^{-4} f^{7/6} \text{ nepers},$$

and for the excess integrated absorption in the reflected ray

$$1 \cdot 34 \times 10^{-2} f^{7/6} \text{ nepers}.$$

As before f is measured in megacycles per second. These values are plotted in Figure 10 (b).

ON THE GALVANOMAGNETIC, THERMOMAGNETIC, AND THERMO-ELECTRIC EFFECTS IN ISOTROPIC METALS AND SEMICONDUCTORS

By E. J. MOORE*

[*Manuscript received December 18, 1957*]

Summary

The calculation of the galvanomagnetic, thermomagnetic, and thermoelectric effects in those isotropic metals and semiconductors which can be represented by the two-band model is simplified and extended to cases not previously treated. General expressions are obtained for effects involving time-independent electric fields, magnetic fields, and thermal gradients. Analytic formulae are given for the isothermal galvanomagnetic effects when a high frequency electric field is applied together with a time-independent magnetic field.

Throughout, emphasis is placed on coefficients which are experimentally or theoretically important. The paper concludes with an appendix on the estimation of parameters occurring in the two-band model of metals and semiconductors.

I. INTRODUCTION

The exact calculation of the galvanomagnetic, thermomagnetic, and thermoelectric effects in isotropic metals or semiconductors is extremely difficult and has not yet been attempted. Approximate treatments have been developed by many authors (e.g. Wilson 1953, Ch. 8; Donovan 1954; Madelung 1954), but even these are complicated. Our object in the present paper is to show how these approximate calculations may be simplified and extended.

The two main stages in all these approximate treatments are:

- (1) the derivation from the Boltzmann transport equation of expressions for the electric and thermal current densities in terms of the applied electric fields, magnetic fields, and thermal gradients;
- (2) the calculation from these equations of formulae for the above effects.

Due to the complex structure of real metals and semiconductors it is necessary to base stage (1) of the calculations on some simplified theoretical model. In past work, the free-electron model and the two-band model have been the two most frequently employed. Since the two-band model (Wilson 1953, pp. 43, 198-9) is the most general one for which the calculations can be carried to completion, we adopt it in the present discussion.

For the most commonly studied effects which are either one-dimensional (e.g. conductivities) or two-dimensional (e.g. Hall effect), two components of each current density must be specified. In the past, one equation has been given for each component (cf. Wilson 1953, p. 219; Donovan 1954; Madelung 1954). Recently, however, Dingle (1956) has shown that, if two imaginaries are used

* Physics Department, University of Western Australia, Nedlands, W.A.

instead of one, it is possible to describe the two electric current components with only one equation. This work, which suggested the present investigation, indicated that a simplification could probably be effected in the calculation of the isothermal galvanomagnetic effects.

Dingle based his treatment on the one-band (free-electron) model and assumed that harmonically varying electric and time-independent magnetic fields were applied. In order to simplify the calculation of the galvanomagnetic, thermomagnetic, and thermoelectric effects for the isotropic two-band model, both the electric and thermal current densities are required for the case when time-independent thermal gradients also exist. To satisfy these requirements it is necessary to extend Dingle's work. This is done in Section II, where two equations (2.9, 2.10) are obtained for the current densities to replace the four customarily introduced in earlier work.

Previously, stage (2) of the calculations has been attempted separately for :

(a) steady electric field effects in (i) metals (e.g. Wilson 1953, Ch. 8) and (ii) semiconductors (e.g. Madelung, loc. cit.) ;

(b) high frequency electric field effects in (i) metals (Donovan, loc. cit. ; Donovan and Sondheimer 1953), and (ii) semiconductors (left untreated).

In the present discussion, we still consider (a) and (b) separately (Sections III, IV) as the effects in each case are of basically different type (cf. Section IV (b)). As most of the important effects in (a) have been calculated previously, our main object here is to simplify these calculations. However, in (b) past work has been confined to metals, and, moreover, the results obtained for the surface resistance and magnetoresistance have only been extracted numerically (Donovan, loc. cit.). Here, our main aims are to extend the calculations to include semiconductors, and to obtain analytic expressions for the principal galvanomagnetic effects.

For both (a) and (b), most previous authors appear to have been primarily interested in either metals or semiconductors, not both, and consequently their results apply to either the one or the other. In the present paper, we derive (Sections III (c), IV (b)) general formulae which may easily be specialized to either case.

II. GENERAL EQUATIONS FOR THE ELECTRIC AND THERMAL CURRENT DENSITIES

As indicated in the introduction we base this first stage of the calculations on the metal or semiconductor model in which there are two conduction bands of standard form (Wilson 1953, pp. 43, 198-9). The band of greater energy is assumed to be normal (i.e. the charge carriers are negative) and that of lesser energy inverted (positive charge carriers). We make the usual assumption that these bands are independent, and therefore we may consider each one separately.

(a) *Normal Band*

The following derivation of the formulae for the current densities broadly follows that given by Wilson (1953, p. 210), but with modifications suggested by Dingle's paper (loc. cit.). We assume that a "universal" time of relaxation

can be defined, and therefore in the presence of harmonically varying electric fields, and time-independent magnetic fields and thermal gradients, the Boltzmann transport equation may be written (in Gaussian units):

$$(-2\pi\varepsilon/\hbar)(\mathbf{E} + c^{-1}\mathbf{v} \wedge \mathbf{H}) \cdot \text{grad}_{\mathbf{k}} f_n + \mathbf{v} \cdot \text{grad}_{\mathbf{r}} f_n = -(f_n - f_{0n})[(1 + i\omega\tau_n)/\tau_n], \quad \dots\dots\dots (2.1)$$

where $i^2 = -1$, ε is the numerical electronic charge, \mathbf{v} , \mathbf{r} , and \mathbf{k} are respectively the velocity, position vector, and wave vector of the electrons, τ_n is the relaxation time, f_n and f_{0n} are the equilibrium distribution functions of the electrons in the presence and absence respectively of applied fields, \mathbf{E} is the electric field which is proportional to $e^{i\omega t}$, and \mathbf{H} is the steady magnetic field. We also make the usual assumption that f_{0n} is the Fermi-Dirac function $\{\exp [(E_n - \zeta_n)/kT] + 1\}^{-1}$.

To terms linear in the electric fields and thermal gradients, this equation has the solution

$$f_n - f_{0n} = -\mathbf{k} \cdot \mathbf{c}_n(E_n)(\partial f_{0n}/\partial E_n),$$

where $E_n = \frac{1}{2}\hbar^2 |\mathbf{k}|^2/m_n$, m_n being the effective mass of the electrons, and $\mathbf{c}_n = (c_{1n}, c_{2n}, c_{3n})$.

The analysis then proceeds as in Wilson (*loc. cit.*), and for the special case of fields given by:

$$\begin{aligned} \text{Electric field} \quad \mathbf{E} &= (E_x, E_y, 0), \\ \text{Magnetic field} \quad \mathbf{H} &= (0, 0, H), \\ \text{Thermal gradient} \quad \mathbf{K} &= (K_x, K_y, 0) = (\partial T/\partial x, \partial T/\partial y, 0), \end{aligned}$$

we obtain (dropping the subscript n)*

$$\left. \begin{aligned} \frac{2\pi m}{\hbar\tau}(1 + i\omega\tau)c_1 - \frac{2\pi m\Omega}{h}c_2 &= -\varepsilon E_x - \left[T \frac{\partial}{\partial T} \left(\frac{\zeta}{T} \right) + \frac{E}{T} \right] K_x, \\ \frac{2\pi m\Omega}{h}c_1 + \frac{2\pi m}{\hbar\tau}(1 + i\omega\tau)c_2 &= -\varepsilon E_y - \left[T \frac{\partial}{\partial T} \left(\frac{\zeta}{T} \right) + \frac{E}{T} \right] K_y, \end{aligned} \right\} \dots (2.2)$$

where $\Omega = (-\varepsilon)H/mc$ is the circular frequency of precessional motion of an electron.

Now, if we define $\mathbf{C} = c_1 + jc_2$, $\mathbf{E} = E_x + jE_y$, and $\mathbf{K} = K_x + jK_y$, with $j^2 = -1$ (but $ij \neq -1$), equations (2.2) reduce to

$$\mathbf{C} = -\frac{h}{2\pi m} \left[\frac{\tau(\varepsilon\mathbf{E} + \beta\mathbf{K})}{1 + (i\omega + j\Omega)\tau} \right], \quad \dots\dots\dots (2.3)$$

where

$$\beta = T \frac{\partial}{\partial T} \left(\frac{\zeta}{T} \right) + \frac{E}{T}.$$

* It is implicitly assumed that the Fermi energy ζ is a function of temperature only, i.e. that the substance is homogeneous.

(i) *Electric Current Density*.—The electric current density $\mathbf{J}=(J_x, J_y, 0)$ has components :

$$J_x = \frac{32\sqrt{2}\pi^2\epsilon m^{3/2}}{3h^4} \int_0^\infty E^{3/2} c_1 \frac{\partial f_0}{\partial E} dE, \quad J_y = \frac{32\sqrt{2}\pi^2\epsilon m^{3/2}}{3h^4} \int_0^\infty E^{3/2} c_2 \frac{\partial f_0}{\partial E} dE.$$

We introduce $\mathbf{J}=\mathbf{J}_x+\mathbf{j}\mathbf{J}_y$, and obtain

$$\mathbf{J} = \frac{32\sqrt{2}\pi^2\epsilon m^{3/2}}{3h^4} \int_0^\infty E^{3/2} \mathbf{C} \frac{\partial f_0}{\partial E} dE,$$

that is,

$$\mathbf{J}_n = -Mm_n^{\frac{1}{2}} \int_0^\infty \frac{\tau_n(\epsilon\mathbf{E}+\beta_n\mathbf{K})}{1+(i\omega+\mathbf{j}\Omega_n)\tau_n} E_n^{3/2} \frac{\partial f_{0n}}{\partial E_n} dE_n, \quad \dots\dots (2.4)$$

where $M=16\sqrt{2}\pi\epsilon/3h^3$.

(ii) *Thermal Current Density*.—The thermal current density $\mathbf{W}=(W_x, W_y, 0)$ has components :

$$\left. \begin{aligned} W_x &= -\frac{32\sqrt{2}\pi^2 m^{3/2}}{3h^4} \int_0^\infty E' E^{3/2} c_1 \frac{\partial f_0}{\partial E} dE, \\ W_y &= -\frac{32\sqrt{2}\pi^2 m^{3/2}}{3h^4} \int_0^\infty E' E^{3/2} c_2 \frac{\partial f_0}{\partial E} dE, \end{aligned} \right\} \dots\dots (2.5)$$

where $E'=A_n+E$ and A_n is the energy of the base of the normal band when the energy at the metal or semiconductor surface is defined to be zero. In the past, the energy zero has generally been taken either at the bottom of the normal band (e.g. Wilson 1953, p. 219) or as the mean of the energies of the bottom of the normal and the top of the inverted bands (e.g. Madelung 1954). Here, we have defined the energy zero as above, since all experimental measurements are made at the metal or semiconductor surface. However, the choice of energy zero only affects the results when both the electric current density and the thermal gradient are simultaneously non-zero.

From (2.3) and (2.5) we find that

$$\mathbf{W}_n = \mathbf{W}_x + \mathbf{j}\mathbf{W}_y = -\frac{A_n}{\epsilon} \mathbf{J}_n + \frac{Mm_n^{\frac{1}{2}}}{\epsilon} \int_0^\infty \frac{\tau_n(\epsilon\mathbf{E}+\beta_n\mathbf{K})}{1+(i\omega+\mathbf{j}\Omega_n)\tau_n} E_n^{5/2} \frac{\partial f_{0n}}{\partial E_n} dE_n. \quad \dots (2.6)$$

(b) *Inverted Band*

The current densities for the inverted band are derived by a method similar to that in Section II (a) (see Wilson 1953, p. 211). We find

$$\mathbf{J}_p = -Mm_p^{\frac{1}{2}} \int_0^\infty \frac{\tau_p(\epsilon\mathbf{E}-\beta_p\mathbf{K})}{1+(i\omega+\mathbf{j}\Omega_p)\tau_p} E_p^{3/2} \frac{\partial f_{0p}}{\partial E_p} dE_p, \quad \dots\dots\dots (2.7)$$

$$\mathbf{W}_p = -\frac{A_p}{\epsilon} \mathbf{J}_p - \frac{Mm_p^{\frac{1}{2}}}{\epsilon} \int_0^\infty \frac{\tau_p(\epsilon\mathbf{E}-\beta_p\mathbf{K})}{1+(i\omega+\mathbf{j}\Omega_p)\tau_p} E_p^{5/2} \frac{\partial f_{0p}}{\partial E_p} dE_p, \quad \dots\dots (2.8)$$

where $\Omega_p=\epsilon H/m_p c$ and A_p is the energy of the top of the inverted band.

(c) *Total Current Densities*

As we have assumed that the two bands are independent, the total current densities are the sums of the current densities for each band. We therefore obtain from (2.4) and (2.7)

$$\mathbf{J} = \mathbf{J}_n + \mathbf{J}_p = -M \int_0^\infty \left[\frac{m_n^{\frac{1}{2}} \tau_n (\varepsilon \mathbf{E} + \beta_n \mathbf{K}) E_n^{3/2}}{1 + (i\omega + j\Omega_n) \tau_n} \frac{\partial f_{0n}}{\partial E_n} dE_n + \frac{m_p^{\frac{1}{2}} \tau_p (\varepsilon \mathbf{E} - \beta_p \mathbf{K}) E_p^{3/2}}{1 + (i\omega + j\Omega_p) \tau_p} \frac{\partial f_{0p}}{\partial E_p} dE_p \right], \quad (2.9)$$

and from (2.6) and (2.8)

$$\mathbf{W} = -\kappa_L \mathbf{K} + \mathbf{W}_n + \mathbf{W}_p = -\kappa_L \mathbf{K} - \varepsilon^{-1} (A_n \mathbf{J}_n + A_p \mathbf{J}_p) + \frac{M}{\varepsilon} \int_0^\infty \left[\frac{m_n^{\frac{1}{2}} \tau_n (\varepsilon \mathbf{E} + \beta_n \mathbf{K}) E_n^{5/2}}{1 + (i\omega + j\Omega_n) \tau_n} \frac{\partial f_{0n}}{\partial E_n} dE_n - \frac{m_p^{\frac{1}{2}} \tau_p (\varepsilon \mathbf{E} - \beta_p \mathbf{K}) E_p^{5/2}}{1 + (i\omega + j\Omega_p) \tau_p} \frac{\partial f_{0p}}{\partial E_p} dE_p \right], \quad (2.10)$$

where κ_L is the thermal conductivity of the lattice.

We are now in a position to commence the second stage of the calculations, which, as indicated in the introduction, must be carried out separately for the two special cases of steady, and high frequency, electric fields. As the time-independent case is the more important it is considered first.

III. STEADY ELECTRIC FIELD EFFECTS

(a) *Introduction*

All the steady electric field effects are defined in terms of the components of the fields and the current densities. Since the x and y components of the current densities are the real and imaginary parts respectively of (2.9) and (2.10), it is now advisable to rationalize these equations (with $\omega=0$). This is a simple process which yields

$$\mathbf{J} = [\Phi_{EL}(\Omega) + jH\Phi_{ET}(\Omega)] \mathbf{E} + [\Phi_{KL}(\Omega) + jH\Phi_{KT}(\Omega)] \mathbf{K}, \quad (3.1)$$

$$\mathbf{W} = [\Psi_{EL}(\Omega) + jH\Psi_{ET}(\Omega)] \mathbf{E} + [\Psi_{KL}(\Omega) + jH\Psi_{KT}(\Omega)] \mathbf{K}, \quad (3.2)$$

where

$$\Phi_{EL}(\Omega) = \varepsilon [I_{11n}(\Omega_n) + I_{11p}(\Omega_p)], \quad (3.3)$$

$$\Phi_{ET}(\Omega) = \varepsilon^2 c^{-1} [m_n^{-1} I_{21n}(\Omega_n) - m_p^{-1} I_{21p}(\Omega_p)], \quad (3.4)$$

$$\Phi_{KL}(\Omega) = T \left[I_{11n}(\Omega_n) \frac{\partial}{\partial T} \left(\frac{\zeta_n}{T} \right) - I_{11p}(\Omega_p) \frac{\partial}{\partial T} \left(\frac{\zeta_p}{T} \right) \right] + \frac{1}{T} [I_{12n}(\Omega_n) - I_{12p}(\Omega_p)], \quad (3.5)$$

$$\Phi_{KT}(\Omega) = \varepsilon c^{-1} \left\{ T \left[m_n^{-1} I_{21n}(\Omega_n) \frac{\partial}{\partial T} \left(\frac{\zeta_n}{T} \right) + m_p^{-1} I_{21p}(\Omega_p) \frac{\partial}{\partial T} \left(\frac{\zeta_p}{T} \right) \right] + T^{-1} [m_n^{-1} I_{22n}(\Omega_n) + m_p^{-1} I_{22p}(\Omega_p)] \right\}, \quad (3.6)$$

$$\Psi_{EL}(\Omega) = -\varepsilon^{-1} [A_n \Phi_{ELn}(\Omega_n) + A_p \Phi_{ELp}(\Omega_p)] - [I_{12n}(\Omega_n) - I_{12p}(\Omega_p)], \quad (3.7)$$

$$\Psi_{ET}(\Omega) = -\varepsilon^{-1} [A_n \Phi_{ETn}(\Omega_n) + A_p \Phi_{ETp}(\Omega_p)] - \{ \varepsilon c^{-1} [m_n^{-1} I_{22n}(\Omega_n) + m_p^{-1} I_{22p}(\Omega_p)] \}, \quad (3.8)$$

$$\Psi_{KL}(\Omega) = -\kappa_L - \varepsilon^{-1} [A_n \Phi_{KLn}(\Omega_n) + A_p \Phi_{KLp}(\Omega_p)] \\ - \varepsilon^{-1} \left\{ T \left[I_{12n}(\Omega_n) \frac{\partial}{\partial T} \left(\frac{\zeta_n}{T} \right) + I_{12p}(\Omega_p) \frac{\partial}{\partial T} \left(\frac{\zeta_p}{T} \right) \right] + \frac{1}{T} [I_{13n}(\Omega_n) + I_{13p}(\Omega_p)] \right\}, \\ \dots \dots \dots (3.9)$$

$$\Psi_{KT}(\Omega) = -\varepsilon^{-1} [A_n \Phi_{KTn}(\Omega_n) + A_p \Phi_{KTp}(\Omega_p)] - c^{-1} \left\{ T \left[m_n^{-1} I_{22n}(\Omega_n) \frac{\partial}{\partial T} \left(\frac{\zeta_n}{T} \right) \right. \right. \\ \left. \left. - m_p^{-1} I_{22p}(\Omega_p) \frac{\partial}{\partial T} \left(\frac{\zeta_p}{T} \right) \right] + T^{-1} [m_n^{-1} I_{23n}(\Omega_n) - m_p^{-1} I_{23p}(\Omega_p)] \right\}, \dots (3.10)$$

and, for instance,

$$I_{stn}(x_n) = M m_n^{\frac{1}{2}} L_{stn}(x_n), \\ L_{stn}(x_n) = - \int_0^\infty \frac{\tau_n^2 E_n^{\frac{1}{2} + \frac{1}{2}}}{1 + x_n^2 \tau_n^2} \frac{\partial f_{0n}}{\partial E_n} dE_n.$$

In order to express the Φ 's and Ψ 's entirely in terms of elementary or tabulated functions of the parameters, it is only necessary to evaluate the integral $L_{st}(x)$. This can be readily achieved in the following special cases (see Appendix I):

(1) The electrons (or holes) form a strongly degenerate system—at normal temperatures this applies to all metals and semi-metals, and to some semi-conductors.

(2) The electrons (or holes) form a weakly degenerate system—most semi-conductors are in this group. For the more important scattering mechanisms (thermal, neutral-impurity, and ionic-impurity) the above integral can be expressed in terms of functions tabulated by Johnson and Whitesell (1953), Dingle, Arndt, and Roy (1957*a*, 1957*b*, 1957*c*), and Beer, Armstrong, and Greenberg (1957).

We are now in a position to derive formulae for the coefficients which describe the various steady electric field effects. In general, previous authors (e.g. Wilson 1953, Ch. 8; Bass and Tsidil'kovskii 1956; Beer, Armstrong, and Greenberg 1957) have expressed these coefficients directly in terms of the metal or semiconductor parameters, i.e. as functions of the expressions equivalent to equations (3.3)–(3.10). In the present paper we give all our formulae in terms of the Φ 's and Ψ 's (cf. Madelung 1954). The advantages of introducing these eight functions are:

(a) The calculations are simplified and condensed.

(b) The final expressions obtained for the coefficients are simpler, and hence any relationships between the various effects are more easily deduced.

(c) Once the values of these functions are determined, either theoretically or experimentally, any essentially one- or two-dimensional coefficient may be readily evaluated. Possible methods of measuring these functions experimentally are indicated in the subsequent work (Section III (c)).

(d) From Φ_{EL} , Φ_{ET} , Ψ_{EL} , and Ψ_{ET} , the values of the principal two-band model metal and semiconductor parameters may be estimated (see Appendix I).

For metals, the parameters A_n , A_p , τ_n , τ_p , m_n , m_p , ζ_n , ζ_p , and hence the number of free electrons (n) and holes (p) per unit volume may be determined. As far as the author can ascertain, no method has previously been given for the evaluation of all the above metal parameters from measurements of the galvanomagnetic, thermomagnetic, and thermoelectric effects alone. Moreover, earlier authors have made simplifying assumptions (e.g. $n=p$, $\tau_n=\tau_p$; cf. Wilson (1953, p. 217)), thus reducing the value of the two-band model.

It is interesting to note that, as all the essentially one- or two-dimensional effects can be expressed in terms of these eight functions, only a maximum of eight of these effects may be considered to be independent. It can easily be shown that this applies to anisotropic as well as to isotropic materials.

(b) Notation

As the experimental conditions under which each effect is measured may vary considerably, it is an advantage to adopt some standard notation to indicate the combination under consideration. In the present paper we indicate the conditions by the superscripts in Table 1.

TABLE 1
NOTATION EMPLOYED FOR INDICATING EXPERIMENTAL
CONDITIONS

Superscript	Condition
E	$E_x=0$
e	$E_y=0$
J	$J_x=0$
j	$J_y=0$
K	$K_x=0$
k	$K_y=0$
W	$W_x=0$
w	$W_y=0$
0	$H=0$

The x direction is taken as longitudinal and the y direction as transverse.

For instance, the isothermal electrical conductivity for non-zero magnetic fields and zero transverse electric current is denoted by σ^{jKk} .

(c) General Formulae* for the Steady Electric Field Effects

Since there are many coefficients describing the various effects, it is impossible to include them all in a paper of reasonable length. We therefore consider only those which satisfy one or other of the following criteria:

(a) the coefficient is experimentally important, or

(b) it is an extremely simple function of the Φ 's and Ψ 's, and is experimentally measurable.

* Unless stated otherwise, the formulae in this subsection are derived from equations (3.1) and/or (3.2).

The reason for employing (a) is obvious. Condition (b) is adopted since the coefficients satisfying it provide the simplest method of determining the Φ 's and Ψ 's, and hence of evaluating the parameters of a metal or semiconductor [Section III (a) (d)].

(i) *Conductivities*.—(1) *Electrical conductivity*.—The electrical conductivity is defined by $\sigma = J_x/E_x$. Only the isothermal conductivities are treated here since they are the more important experimentally and also the simpler theoretically.

The principal cases are :

Zero Magnetic Field

$$\sigma^{K0} = \Phi_{EL}(0), \quad \dots\dots\dots (3.11)$$

Non-zero Magnetic Field

(a) $E_y = 0$

$$\sigma^{eKk} = \Phi_{EL}(\Omega). \quad \dots\dots\dots (3.12)$$

The measurement of this quantity immediately yields experimental values for the function Φ_{EL} , the exact determination of which is essential for reliable parameter estimation (see Appendix I). This conductivity is also required in the derivation of the Corbino magnetoresistance (Section III (c) (ii) (1)).

(b) $J_y = 0$

Although it is not particularly important, this conductivity is included here since it is required for the calculation of the normal magnetoresistance.

We find that

$$\sigma^{jKk} = \Phi_{EL}(\Omega) + [H^2 \Phi_{ET}^2(\Omega)/\Phi_{EL}(\Omega)]. \quad \dots\dots\dots (3.13)$$

(2) *Thermal conductivity*. — The thermal conductivity is defined by $\kappa = -W_x/K_x$. The usual experimental conditions are $\mathbf{J} = 0$ and $W_y = 0$, but in this case the formulae obtained for κ contain products of the Φ 's and Ψ 's. For $\mathbf{E} = 0$ and $K_y = 0$, the expressions derived for κ are much simpler, and therefore they would be expected to yield more precise values for the Φ 's and Ψ 's. However, these latter conditions are more difficult to apply in practice, and little, if any, work has been carried out for them.

We can readily deduce the following equations for these two cases :

Zero Magnetic Field

$$(a) \quad \kappa^{E0} = -\Psi_{KL}(0), \quad \dots\dots\dots (3.14)$$

$$(b) \quad \kappa^{J0} = -\Psi_{KL}(0) + [\Psi_{EL}(0)\Phi_{KL}(0)/\Phi_{EL}(0)]. \quad \dots\dots (3.15)$$

Non-zero Magnetic Field

$$(a) \quad \kappa^{Eek} = -\Psi_{KL}(\Omega). \quad \dots\dots\dots (3.16)$$

The measurement of κ^{Eek} would therefore yield values of $\Psi_{KL}(\Omega)$.

$$(b) \quad \chi^{Jjw} = - \frac{\begin{vmatrix} \Phi_{EL} & -H\Phi_{ET} & \Phi_{KL} & -H\Phi_{KT} \\ H\Phi_{ET} & \Phi_{EL} & H\Phi_{KT} & \Phi_{KL} \\ \Psi_{EL} & -H\Psi_{ET} & \Psi_{KL} & -H\Psi_{KT} \\ H\Psi_{ET} & \Psi_{EL} & H\Psi_{KT} & \Psi_{KL} \end{vmatrix}}{\begin{vmatrix} \Phi_{EL} & -H\Phi_{ET} & -H\Phi_{KT} \\ H\Phi_{ET} & \Phi_{EL} & \Phi_{KL} \\ H\Psi_{ET} & \Psi_{EL} & \Psi_{KL} \end{vmatrix}}. \quad (3.17)$$

(ii) *Galvanomagnetic Effects*.—(1) *Corbino magnetoresistance*.—The Corbino magnetoresistive effect is described by the coefficient $M_{\rho e} = (\rho^e - \rho^0)/\rho^0$, where ρ is the resistivity. For the more important isothermal conditions we obtain from (3.11) and (3.12)

$$M_{\rho e}^{Kk} = [\Phi_{EL}(0)/\Phi_{EL}(\Omega)] - 1. \quad (3.18)$$

(2) *Normal magnetoresistance*.—The normal magnetoresistance $M_{\rho j}^{Kk}$ is defined by $M_{\rho j}^{Kk} = (\rho^{jKk} - \rho^{K0})/\rho^{K0}$. Equations (3.11) and (3.13) yield

$$M_{\rho j}^{Kk} = - \frac{H^2 \Phi_{ET}^2(\Omega) + \Phi_{EL}(\Omega)[\Phi_{EL}(\Omega) - \Phi_{EL}(0)]}{\Phi_{EL}^2(\Omega) + H^2 \Phi_{ET}^2(\Omega)}. \quad (3.19)$$

(3) *Hall coefficient*.—The Hall coefficient is $R = E_y/HJ_x$; $J_y = 0$. For the more common isothermal conditions we easily obtain

$$R^{Kk} = - \frac{\Phi_{ET}(\Omega)}{\Phi_{EL}^2(\Omega) + H^2 \Phi_{ET}^2(\Omega)}. \quad (3.20)$$

A simpler closely related effect is defined by

$$S = E_y/H E_x; \quad (J_y = 0) = (\tan \theta)/H,$$

where θ is the Hall angle. Under isothermal conditions we then find that

$$S^{Kk} = -\Phi_{ET}(\Omega)/\Phi_{EL}(\Omega). \quad (3.21)$$

As $\Phi_{EL}(\Omega) = \sigma^{eKk}$, this coefficient gives experimental values for $\Phi_{ET}(\Omega)$, a function which is extremely important for the estimation of parameters (see Appendix I).

(4) *Etingshausen coefficient*.—The Etingshausen coefficient A_E is defined by $A_E = -K_y/HJ_x$; $W_y = 0$. The usual experimental conditions are $J_y = W_x = 0$, in which case

$$A_E^{jW} = - \frac{H \begin{vmatrix} H\Phi_{ET} & \Phi_{EL} & H\Phi_{KT} \\ \Psi_{EL} & -H\Psi_{ET} & \Psi_{KL} \\ H\Psi_{ET} & \Psi_{EL} & H\Psi_{KT} \end{vmatrix}}{\begin{vmatrix} \Phi_{EL} & -H\Phi_{ET} & \Phi_{KL} & -H\Phi_{KT} \\ H\Phi_{ET} & \Phi_{EL} & H\Phi_{KT} & \Phi_{KL} \\ \Psi_{EL} & -H\Psi_{ET} & \Psi_{KL} & -H\Psi_{KT} \\ H\Psi_{ET} & \Psi_{EL} & H\Psi_{KT} & \Psi_{KL} \end{vmatrix}}. \quad (3.22)$$

A more useful coefficient, closely related to the above, is $B_E = -K_y/HE_x$; $W_y=0$. For the conditions $E_y=K_x=0$, it reduces to the extremely simple form

$$B_E^{eK} = \Psi_{ET}(\Omega)/\Psi_{KL}(\Omega). \quad \dots\dots\dots (3.23)$$

Unfortunately, however, it possesses the great practical disadvantage of requiring mixed thermal conditions. If instead of $K_x=0$, we apply the restriction $W_x=0$, the coefficient so defined is more easily obtained experimentally. In this case

$$B_E^{eW} = \frac{\Psi_{KL}\Psi_{ET} - \Psi_{EL}\Psi_{KT}}{\Psi_{KL}^2 + H^2\Psi_{KT}^2}. \quad \dots\dots\dots (3.24)$$

Equation (3.24) provides an experimental method for determining the function Ψ_{ET} , the exact evaluation of which is important for estimating both metal and semiconductor parameters.

(iii) *Thermomagnetic Effects*.—(1) *Magneto thermal conductivity*.—This effect is described by the coefficient $M_x = (\kappa - \kappa^0)/\kappa^0$, which is usually measured subject to the restriction $\mathbf{J}=0$. However, in this case, M_x involves fourth order determinants (see Section III (c) (i) (2)), and therefore it has few theoretical applications. For zero electric field conditions, M_x is much simpler. We find

$$M_x^{Eek} = [\Psi_{KL}(\Omega)/\Psi_{KL}(0)] - 1, \quad \dots\dots\dots (3.25)$$

$$M_x^{Eeo} = \frac{\Psi_{KL}(\Omega) - \Psi_{KL}(0)}{\Psi_{KL}(0)} + \frac{H^2\Psi_{KT}^2(\Omega)}{\Psi_{KL}(\Omega)\Psi_{KL}(0)}. \quad \dots\dots (3.26)$$

(2) *Righi-Leduc coefficient*.—The Righi-Leduc coefficient is $B_{RL} = K_y/HK_x$; $W_y=0$, and is therefore the thermomagnetic analogue of S (Section III (c) (ii) (3)). It is generally measured under conditions of zero electric current, but theoretically, zero electric field conditions possess a distinct advantage. However, experimentally, the reverse is true, and it is doubtful if B_{RL} has yet been determined for these latter conditions. We obtain

$$B_{RL}^{Ee} = -\Psi_{KT}(\Omega)/\Psi_{KL}(\Omega), \quad \dots\dots\dots (3.27)$$

Measurements of this quantity would therefore yield values of $\Psi_{KT}(\Omega)$.

$$B_{RL}^{Jl} = -\frac{H}{H} \begin{vmatrix} \Phi_{EL} & -H\Phi_{ET} & \Phi_{KL} \\ H\Phi_{ET} & \Phi_{EL} & H\Phi_{KT} \\ H\Psi_{ET} & \Psi_{EL} & H\Psi_{KT} \end{vmatrix} \cdot \dots\dots (3.28)$$

$$\begin{vmatrix} \Phi_{EL} & -H\Phi_{ET} & -H\Phi_{KT} \\ H\Phi_{ET} & \Phi_{EL} & \Phi_{KL} \\ H\Psi_{ET} & \Psi_{EL} & \Psi_{KL} \end{vmatrix}$$

(3) *Ettingshausen-Nernst coefficient*.—This coefficient is defined by $B_{EN} = -E_y/HK_x$; $J_y=0$, and is thus the thermomagnetic analogue of B_E (Section III (c) (ii) (4)). In practice, B_{EN} is usually determined subject to the

restrictions $J_x=0$ and either $K_y=0$ or $W_y=0$. However, theoretically, the simplest and therefore the most useful conditions are $E_x=0$ and $K_y=0$. We find

$$B_{EN}^{Ek} = \Phi_{KT}(\Omega)/\Phi_{EL}(\Omega), \quad \dots\dots\dots (3.29)$$

A measurement of this quantity would thus determine $\Phi_{KT}(\Omega)$.

$$B_{EN}^{Jk} = \frac{\Phi_{EL}\Phi_{KT} - \Phi_{ET}\Phi_{KL}}{\Phi_{EL}^2 + H^2\Phi_{ET}^2}. \quad \dots\dots\dots (3.30)$$

As values of Φ_{EL} , Φ_{ET} , and Φ_{KL} are obtainable, equation (3.30) provides an alternative method of evaluating $\Phi_{KT}(\Omega)$.

(iv) *Thermoelectric Effects*.—In the present discussion we consider only those effects which apply to a single metal or semiconductor. Thus, for example, the Peltier effect is not treated.

(1) *Thomson coefficient*.—The Thomson coefficient μ is defined as the coefficient of $-K_x J_x$ in the equation $Q = \mathbf{E} \cdot \mathbf{J} - \text{div } \mathbf{W}$, which gives the rate of energy production per unit volume of a substance in which an electric field, electric current, and thermal current are present. For zero magnetic fields we readily obtain

$$\mu^0 = [\Phi_{KL}(0)/\Phi_{EL}(0)] + (\partial/\partial T)[\Psi_{EL}(0)/\Phi_{EL}(0)]. \quad \dots (3.31)$$

It is also easy to calculate μ for a non-zero magnetic field, but as this case does not appear to be of any experimental or theoretical importance it is not treated here.

(2) *Other coefficients*.—Other thermoelectric effects which may become important due to their theoretical simplicity are as follows:

$$(a)^* \quad C^{ek} = E_x/K_x; \quad J_x=0, \quad (E_y=K_y=0) = -\Phi_{KL}/\Phi_{EL}. \quad \dots (3.32)$$

A determination of this quantity therefore enables Φ_{KL} to be obtained.

$$(b) \quad D_W^{ek} = W_x/E_x; \quad K_x=0, \quad (E_y=K_y=0) = \Psi_{EL}(\Omega). \quad \dots\dots (3.33)$$

$$(c) \quad D_K^{ek} = K_x/E_x; \quad W_x=0, \quad (E_y=K_y=0) = -\Psi_{EL}/\Psi_{KL}. \quad \dots (3.34)$$

A measurement of either (b) or (c) would thus yield the value of $\Psi_{EL}(\Omega)$. The evaluation of this function is important for estimating both metal and semiconductor parameters (see Appendix I). Unfortunately, neither of the last two effects appears to have been investigated experimentally up to the present time.

IV. HIGH FREQUENCY ELECTRIC FIELD EFFECTS

(a) Introduction

For alternating electric fields only the isothermal galvanomagnetic effects are experimentally important. Under isothermal conditions the equation (2.9) for the electric current density becomes

$$\mathbf{J} = -M\epsilon \int_0^\infty \left[\frac{m_n^{\frac{1}{2}}\tau_n E_n^{3/2}}{1 + (i\omega + j\Omega_n)\tau_n} \frac{\partial f_{0n}}{\partial E_n} dE_n + \frac{m_p^{\frac{1}{2}}\tau_p E_p^{3/2}}{1 + (i\omega + j\Omega_p)\tau_p} \frac{\partial f_{0p}}{\partial E_p} dE_p \right] \mathbf{E} \dots (4.1)$$

* The name "Beer-Willardson" effect has been suggested for this by Armstrong and Greenberg (Beer, Armstrong, and Greenberg 1957).

As in Section III, it is now desirable to separate the real and imaginary parts of this expression. By substituting

$$\frac{2}{1 + (i\omega + j\Omega)\tau} = \left[\frac{1}{1 + (\omega + \Omega)^2\tau^2} + \frac{1}{1 + (\omega - \Omega)^2\tau^2} \right] - i \left[\frac{(\omega + \Omega)\tau}{1 + (\omega + \Omega)^2\tau^2} + \frac{(\omega - \Omega)\tau}{1 + (\omega - \Omega)^2\tau^2} \right] - j \left[\frac{(\omega + \Omega)\tau}{1 + (\omega + \Omega)^2\tau^2} + \frac{(\Omega - \omega)\tau}{1 + (\omega - \Omega)^2\tau^2} \right] - ij \left[\frac{1}{1 + (\omega + \Omega)^2\tau^2} - \frac{1}{1 + (\omega - \Omega)^2\tau^2} \right]$$

in (4.1) we can readily derive the following rationalized equation

$$\mathbf{J} = \{ [\Pi_{ELR}(\Omega, \omega) + i\Pi_{ELI}(\Omega, \omega)] + jH[\Pi_{ETR}(\Omega, \omega) + i\Pi_{ETI}(\Omega, \omega)] \} \mathbf{E}, \quad \dots \quad (4.2)$$

where

$$\Pi_{ELR}(\Omega, \omega) = \frac{1}{2}\varepsilon[I_{11n}(\omega + \Omega_n) + I_{11n}(\omega - \Omega_n) + I_{11p}(\omega + \Omega_p) + I_{11p}(\omega - \Omega_p)], \quad \dots \dots \dots (4.3)$$

$$\begin{aligned} \Pi_{ELI}(\Omega, \omega) = & -\frac{1}{2}\varepsilon\{\omega[I_{21n}(\omega + \Omega_n) + I_{21n}(\omega - \Omega_n) + I_{21p}(\omega + \Omega_p) + I_{21p}(\omega - \Omega_p)] \\ & + \Omega_n[I_{21n}(\omega + \Omega_n) - I_{21n}(\omega - \Omega_n)] + \Omega_p[I_{21p}(\omega + \Omega_p) - I_{21p}(\omega - \Omega_p)]\}, \\ & \dots \dots \dots (4.4) \end{aligned}$$

$$\begin{aligned} \Pi_{ETR}(\Omega, \omega) = & \frac{1}{2}\varepsilon^2\sigma^{-1}\{m_n^{-1}[I_{21n}(\omega + \Omega_n) + I_{21n}(\omega - \Omega_n)] \\ & - m_p^{-1}[I_{21p}(\omega + \Omega_p) + I_{21p}(\omega - \Omega_p)] \\ & - (c\omega/\varepsilon H)[I_{21n}(\omega + \Omega_n) - I_{21n}(\omega - \Omega_n) + I_{21p}(\omega + \Omega_p) - I_{21p}(\omega - \Omega_p)]\}, \\ & \dots \dots \dots (4.5) \end{aligned}$$

$$\Pi_{ETI}(\Omega, \omega) = -\frac{1}{2}\varepsilon H^{-1}[I_{11n}(\omega + \Omega_n) - I_{11n}(\omega - \Omega_n) + I_{11p}(\omega + \Omega_p) - I_{11p}(\omega - \Omega_p)], \quad \dots \dots \dots (4.6)$$

and, as in Section III,

$$I_{stn}(x_n) = -Mm_n^{\frac{1}{2}} \int_0^\infty \frac{\tau_n^s E_n^{t+\frac{1}{2}}}{1 + x_n^2 \tau_n^2} \frac{\partial f_{0n}}{\partial E_n} dE_n.$$

(b) General Formulae for the High Frequency Electric Field Effects

Before attempting to derive any formulae we must first distinguish between two basically different types of effects.

Point Effects.—These are defined in terms of the currents and fields at a particular point in the metal or semiconductor, and are therefore independent of the geometry of the specimen. In general, they are not directly measurable, i.e. the fields and currents appearing in the definition cannot be directly determined. The conductivity is an example of a point effect.

Surface Effects.—In these it is necessary to take into account the variation of the electric field with distance from the surface of the sample, and therefore the geometry of the metal or semiconductor is important. In general, they are directly measurable. Examples of surface effects are the surface impedance and the surface magnetoresistance.

(i) *Point Effects*.—These effects are defined in a similar manner to those in the steady electric field case, and the expressions obtained for them are formally the same, but with $\Pi_{ELR} + i\Pi_{ELI}$ replacing Φ_{EL} , and $\Pi_{ETR} + i\Pi_{ETI}$ replacing Φ_{ET} .

(1) *Electrical conductivity*.—As in Section III (c) (i) (1), the electrical conductivity is defined by $\sigma = J_x/E_x$. The only experimental conditions of any importance are:

Zero Magnetic Field

From (4.2) we obtain

$$\sigma^0 = [\Pi_{ELR}^2(0, \omega) + \Pi_{ELI}^2(0, \omega)]^{\frac{1}{2}} \exp \{i \tan^{-1} [\Pi_{ELI}(0, \omega) / \Pi_{ELR}(0, \omega)]\}. \quad \dots\dots\dots (4.7)$$

Non-zero Magnetic Field

(a) $E_y = 0$

$$\sigma^e = [\Pi_{ELR}^2(\Omega, \omega) + \Pi_{ELI}^2(\Omega, \omega)]^{\frac{1}{2}} \exp \{i \tan^{-1} [\Pi_{ELI}(\Omega, \omega) / \Pi_{ELR}(\Omega, \omega)]\}, \quad \dots\dots\dots (4.8)$$

(b) $J_y = 0$

$$\sigma^j = \sigma^e + \frac{H^2 [\Pi_{ETR}^2 + \Pi_{ETI}^2]}{[\Pi_{ELR}^2 + \Pi_{ELI}^2]^{\frac{1}{2}}} \exp \left\{ i \left[2 \tan^{-1} \left(\frac{\Pi_{ETI}}{\Pi_{ETR}} \right) - \tan^{-1} \left(\frac{\Pi_{ELI}}{\Pi_{ELR}} \right) \right] \right\}. \quad \dots\dots\dots (4.9)$$

(ii) *Surface Effects*.—The following calculations do not take account of the anomalous skin effect (Reuter and Sondheimer 1948). The displacement current is included, although for metals it is only significant when the anomalous terms in the skin effect should also be taken into account. It is considered here since it may be required for semiconductors, due to their lower conductivity, and also because it is as easy to include as to exclude.

It is assumed that the metal or semiconductor specimen occupies the volume defined by $z \geq 0$. The magnetic field is then in the direction of the outward normal from the surface, and the electric fields are parallel to the surface (see Section II).

As indicated in Section IV (b), the electric field is a function of position. With the above assumptions this function satisfies Maxwell's equations, which yield

$$\frac{d^2 \mathbf{E}}{dz^2} = \mathbf{E}'' = \frac{\mu}{c^2} \left(\kappa \frac{\partial^2 \mathbf{E}}{\partial t^2} + 4\pi \frac{\partial \mathbf{J}}{\partial t} \right), \quad \dots\dots\dots (4.10)$$

where μ is the magnetic permeability and κ is the dielectric constant. As it has been implicitly assumed in Section II that $\mu = 1$, we must also make this assumption here, thus excluding ferromagnetic materials from consideration.

On combining (4.2) and (4.10), and remembering that $\partial \mathbf{E} / \partial t = i\omega \mathbf{E}$ (see Section II (a)) we obtain

$$\mathbf{E}'' = (A + jHB)\mathbf{E}, \quad \dots\dots\dots (4.11)$$

where

$$A = \omega c^{-2} \{ -\omega \kappa + 4\pi i [\Pi_{ELR}(\Omega, \omega) + i \Pi_{ELI}(\Omega, \omega)] \}, \quad \dots \quad (4.12)$$

$$B = 4\pi \omega c^{-2} i [\Pi_{ETR}(\Omega, \omega) + i \Pi_{ETI}(\Omega, \omega)]. \quad \dots \quad (4.13)$$

To simplify later work we define here

$$\theta = \arg A = \tan^{-1} \left[-\frac{4\pi \Pi_{ELR}(\Omega, \omega)}{\omega \kappa + 4\pi \Pi_{ELI}(\Omega, \omega)} \right], \quad \dots \quad (4.14)$$

$$\varphi = \arg B = \tan^{-1} \{ -[\Pi_{ETR}(\Omega, \omega) / \Pi_{ETI}(\Omega, \omega)] \}. \quad \dots \quad (4.15)$$

Under the boundary condition that $E(z)$ is always finite, (4.11) possesses the solution

$$\left. \begin{aligned} E_x(z) &= a \exp(-\lambda_1 z) + b \exp(-\lambda_2 z), \\ E_y(z) &= -ja \exp(-\lambda_1 z) + jb \exp(-\lambda_2 z), \end{aligned} \right\} \quad \dots \quad (4.16)$$

where $a = \frac{1}{2}[E_x(0) + jE_y(0)]$, $b = \frac{1}{2}[E_x(0) - jE_y(0)]$, $\lambda_1 = (A + jHB)^{\frac{1}{2}}$, $\lambda_2 = (A - jHB)^{\frac{1}{2}}$, and $-\frac{1}{2}\pi < \arg \lambda_1, \lambda_2 \leq \frac{1}{2}\pi$, i.e. the real parts of both λ_1 and λ_2 are positive. Since we have taken $E \propto e^{i\omega t}$ (Section II) and we require the solution to represent waves travelling in the positive z direction, the imaginary parts of λ_1 and λ_2 must also be positive, i.e. $0 \leq \arg \lambda_1, \lambda_2 \leq \pi$. The final restriction on $\arg \lambda_1, \lambda_2$ is therefore $0 \leq \arg \lambda_1, \lambda_2 \leq \frac{1}{2}\pi$.

We are now in a position to calculate the various surface effects.

(1) *Surface impedance*.—The surface impedance is defined by

$$Z(\omega, H) = E_x(0) \left/ \int_0^\infty J_x(z) dz \right. = -(4\pi \omega c^{-2}) i E_x(0) / E'_x(0). \quad \dots \quad (4.17)$$

For non-zero magnetic fields, one more condition is necessary. The most convenient restrictions are either, (i) $E_y(0) = 0$, or (ii) $\int_0^\infty J_y(z) dz = 0$, i.e. $E'_y(0) = 0$. As (ii) appears to be the more common condition (Donovan 1954) it is assumed here. It is also required in the calculation of the normal magnetoresistance. Condition (i) would define the high frequency analogue of the Corbino magnetoresistance (Section III (c) (ii) (1)), but no attempt is made to calculate this effect here as it is neither experimentally nor theoretically important.

With restriction (ii), we find from (4.16) and (4.17) that

$$Z(\omega, H) = (4\pi \omega c^{-2}) i (\lambda_1 + \lambda_2) / 2\lambda_1 \lambda_2 \quad \dots \quad (4.18)$$

$$\begin{aligned} &= (2\pi \omega c^{-2}) i [(A + jHB)^{\frac{1}{2}} + (A - jHB)^{\frac{1}{2}}] [A^2 + H^2 B^2]^{-\frac{1}{2}}. \\ &\quad \dots \quad (4.19) \end{aligned}$$

Since A and B are complex numbers, any attempt to separate the real and imaginary parts of (4.19) exactly will result in extremely complicated expressions. However, for weak magnetic fields ($H \ll B \ll A \ll 1$), it is comparatively simple

to expand (4.19) in ascending powers of HB/A and then rationalize term by term. We find after expanding and rationalizing (4.19) that

$$Z(\omega, H) = (4\pi\omega c^{-2}) |A|^{-\frac{1}{2}} \exp(-\frac{1}{2}i\theta) \left\{ 1 - \frac{3}{8} \frac{H^2 |B|^2}{|A|^2} \exp[2i(\varphi - \theta)] \right. \\ \left. + \frac{35}{128} \frac{H^4 |B|^4}{|A|^4} \exp[4i(\varphi - \theta)] + \dots \right\}, \quad \dots\dots\dots (4.20)$$

and

$$Z(\omega, 0) = (4\pi\omega c^{-2}) |A_0|^{-\frac{1}{2}} \exp(-\frac{1}{2}i\theta_0), \quad \dots\dots\dots (4.21)$$

where A_0 and θ_0 are the values of A and θ for $H=0$.

Equation (4.19) may also be expanded in descending powers of HB/A to obtain formal expressions for the strong magnetic field case ($H|B|/|A| > 1$). But for metals at least, this condition implies, in general, that $|\Omega_n \tau_n|$ and/or $|\Omega_p \tau_p| > 1$ and the derivation of the Boltzmann equation then breaks down (Wilson 1953, p. 210). We therefore do not consider this case here.

(2) *Surface resistance and reactance.*—The surface resistance $R(\omega, H)$ and surface reactance $X(\omega, H)$ are defined as the real and imaginary parts respectively of $Z(\omega, H)$. From (4.20) and (4.21) we obtain

$$R(\omega, H) = (4\pi\omega c^{-2}) |A|^{-\frac{1}{2}} \left\{ \sin \frac{1}{2}\theta - \frac{3}{8} \frac{H^2 |B|^2}{|A|^2} \sin [\frac{1}{2}(5\theta - 4\varphi)] \right. \\ \left. + \frac{35}{128} \frac{H^4 |B|^4}{|A|^4} \sin [\frac{1}{2}(9\theta - 8\varphi)] + \dots \right\}, \quad \dots\dots\dots (4.22)$$

$$R(\omega, 0) = (4\pi\omega c^{-2}) |A_0|^{-\frac{1}{2}} \sin \frac{1}{2}\theta_0, \quad \dots\dots\dots (4.23)$$

$$X(\omega, H) = (4\pi\omega c^{-2}) |A|^{-\frac{1}{2}} \left\{ \cos \frac{1}{2}\theta - \frac{3}{8} \frac{H^2 |B|^2}{|A|^2} \cos [\frac{1}{2}(5\theta - 4\varphi)] \right. \\ \left. + \frac{35}{128} \frac{H^4 |B|^4}{|A|^4} \cos [\frac{1}{2}(9\theta - 8\varphi)] + \dots \right\}, \quad \dots\dots\dots (4.24)$$

$$X(\omega, 0) = (4\pi\omega c^{-2}) |A_0|^{-\frac{1}{2}} \cos \frac{1}{2}\theta_0. \quad \dots\dots\dots (4.25)$$

(3) *Surface magnetoresistance.*—The magnetoresistance $M_R(\omega, H)$ is defined by

$$M_R(\omega, H) = [R(\omega, H) - R(\omega, 0)]/R(\omega, 0).$$

The substitution of (4.22) and (4.23) into the above leads to

$$M_R(\omega, H) = \left[\frac{\sin \frac{1}{2}\theta}{\sin \frac{1}{2}\theta_0} \frac{|A_0|^{\frac{1}{2}}}{|A|^{\frac{1}{2}}} - 1 \right] - \frac{3}{8} \frac{H^2 |B|^2}{|A|^{5/2}} \frac{|A_0|^{\frac{1}{2}} \sin [\frac{1}{2}(5\theta - 4\varphi)]}{\sin \frac{1}{2}\theta_0} \\ \times \left\{ 1 - \frac{35}{48} \frac{H^2 |B|^2}{|A|^2} \frac{\sin [\frac{1}{2}(9\theta - 8\varphi)]}{\sin [\frac{1}{2}(5\theta - 4\varphi)]} + \dots \right\}. \quad \dots\dots (4.26)$$

(4) *Hall coefficient*.—Following Donovan and Sondheimer (1953) we define the Hall coefficient $Y(\omega, H)$ by

$$\begin{aligned} Y(\omega, H) &= +E_y(0) \left/ \left[H \int_0^\infty J_x(z) dz \right] \right.; \quad \int_0^\infty J_y(z) dz = 0 \\ &= -(4\pi\omega c^{-2}) i E_y(0) / H E'_x(0); \quad E'_y(0) = 0. \end{aligned}$$

From (4.16) we obtain

$$\begin{aligned} Y(\omega, H) &= \mathbf{j} (4\pi\omega c^{-2}) (\lambda_1 - \lambda_2) / (H \cdot 2\lambda_1\lambda_2) \\ &= \mathbf{j} (2\pi\omega c^{-2}) H^{-1} [(A + \mathbf{j}HB)^{\frac{1}{2}} - (A - \mathbf{j}HB)^{\frac{1}{2}}] [A^2 + H^2B^2]^{-\frac{1}{2}}. \\ &\dots\dots\dots (4.27) \end{aligned}$$

As in Section IV (b) (ii) (1), equation (4.27) can be expanded in ascending powers of HB/A , yielding

$$\begin{aligned} Y(\omega, H) &= -\frac{2\pi\omega}{c^2} \frac{|B|}{|A|^{\frac{3}{2}}} \exp \left[\frac{1}{2} i (2\varphi - 3\theta) \right] \left\{ 1 - \frac{5}{8} \frac{H^2 |B|^2}{|A|^2} \exp [2i(\varphi - \theta)] \right. \\ &\quad \left. + \frac{63}{128} \frac{H^4 |B|^4}{|A|^4} \exp [4i(\varphi - \theta)] - \dots \right\}. \dots\dots\dots (4.28) \end{aligned}$$

V. ACKNOWLEDGMENTS

The author wishes to thank Dr. R. B. Dingle for suggesting this problem and for valuable advice, Mr. J. G. Collins for helpful discussions and criticism, and the Commonwealth Scientific and Industrial Research Organization for the award of a Studentship.

VI. REFERENCES

- BASS, F. G., and TSIDIL'KOVSKII, I. M. (1956).—*J. Exp. Theor. Phys.* **31**: 672; *Soviet Phys., J.E.T.P.* **4**: 565 (1957).
 BEER, A. C., ARMSTRONG, J. A., and GREENBERG, I. N. (1957).—*Phys. Rev.* **107**: 1506.
 DINGLE, R. B. (1955).—*Phil. Mag.* **46**: 831.
 DINGLE, R. B. (1956).—*Physica* **22**: 701.
 DINGLE, R. B., ARNDT, D., and ROY, S. K. (1957a).—*Appl. Sci. Res., Hague B* **6**: 144.
 DINGLE, R. B., ARNDT, D., and ROY, S. K. (1957b).—*Appl. Sci. Res., Hague B* **6**: 155.
 DINGLE, R. B., ARNDT, D., and ROY, S. K. (1957c).—*Appl. Sci. Res., Hague B* **6**: 245.
 DONOVAN, B. (1954).—*Proc. Phys. Soc. Lond. A* **67**: 305.
 DONOVAN, B., and SONDHEIMER, E. H. (1953).—*Proc. Phys. Soc. Lond. A* **66**: 849.
 JOHNSON, V. A., and WHITESELL, W. J. (1953).—*Phys. Rev.* **89**: 941.
 MADELUNG, O. (1954).—*Z. Naturf.* **9a**: 667.
 REUTER, G. E. H., and SONDHEIMER, E. H. (1948).—*Proc. Roy. Soc. A* **195**: 336.
 SHOCKLEY, W. (1950).—"Electrons and Holes in Semiconductors." (D. Van Nostrand: New York.)
 SOMMERFELD, A. (1928).—*Z. Phys.* **47**: 1.
 WILSON, A. H. (1953).—"The Theory of Metals." 2nd Ed. (Cambridge Univ. Press.)

APPENDIX I

Estimation of Two-band Model Parameters

The parameters of a metal or semiconductor may be determined from the experimentally measurable functions Φ_{EL} , Φ_{ET} , Ψ_{EL} , and Ψ_{ET} defined in Section III (a). It is convenient to consider this problem separately for metals and semiconductors since, in general, the above functions assume different forms for each case.

(i) *Metals*.—In metals the electrons and holes form strongly degenerate systems, i.e. Fermi energy $\zeta \gg kT$. Under these conditions, convenient experimental parameters are A_n , A_p , m_n , m_p , τ_n , τ_p , ζ_n , and ζ_p (see Section II). We now proceed to show how these quantities may be estimated from the four functions Φ_{EL} , Φ_{ET} , Ψ_{EL} , and Ψ_{ET} .

All the above functions have been expressed in Section III (a) in terms of the integral

$$L_{st}(x) = - \int_0^\infty \frac{\tau^s E^{t+\frac{1}{2}}}{1+x^2\tau^2} \frac{\partial f_0}{\partial E} dE,$$

which, for $\zeta \gg kT$, may be evaluated by applying the asymptotic expansion first given by Sommerfeld (1928)

$$- \int_0^\infty \varphi(E) \frac{\partial f_0}{\partial E} dE = \varphi(\zeta) + \frac{1}{6}(\pi kT)^2 \frac{\partial^2 \varphi(\zeta)}{\partial \zeta^2} + \dots \quad (1)$$

From (1) we obtain, to first order terms in kT/ζ ,

$$L_{st}(x) = \frac{\tau^s \zeta^{t+\frac{1}{2}}}{1+x^2\tau^2} \quad (2)$$

Substitution of (2) in (3.3), (3.4), (3.7), and (3.9) yields

$$\Phi_{EL}(0) = M\varepsilon(\tau_n m_n^{\frac{1}{2}} \zeta_n^{3/2} + \tau_p m_p^{\frac{1}{2}} \zeta_p^{3/2}) = \sigma_n + \sigma_p, \quad (3)$$

$$\Phi_{EL}(\Omega) = \frac{\sigma_n}{1+\mu_n^2 H^2} + \frac{\sigma_p}{1+\mu_p^2 H^2}, \quad (4)$$

$$\Phi_{ET}(0) = \mu_n \sigma_n - \mu_p \sigma_p, \quad (5)$$

$$\Phi_{ET}(\Omega) = \frac{\mu_n \sigma_n}{1+\mu_n^2 H^2} - \frac{\mu_p \sigma_p}{1+\mu_p^2 H^2}, \quad (6)$$

$$\varepsilon \Psi_{EL}(0) = -(A_n \sigma_n + A_p \sigma_p) - \zeta_n \sigma_n + \zeta_p \sigma_p, \quad (7)$$

$$\varepsilon \Psi_{EL}(\Omega) = - \left(\frac{A_n \sigma_n}{1+\mu_n^2 H^2} + \frac{A_p \sigma_p}{1+\mu_p^2 H^2} \right) - \frac{\zeta_n \sigma_n}{1+\mu_n^2 H^2} + \frac{\zeta_p \sigma_p}{1+\mu_p^2 H^2}, \quad (8)$$

$$\varepsilon \Psi_{ET}(0) = -(A_n \mu_n \sigma_n - A_p \mu_p \sigma_p) - \mu_n \zeta_n \sigma_n - \mu_p \zeta_p \sigma_p, \quad (9)$$

$$\varepsilon \Psi_{ET}(\Omega) = - \left(\frac{A_n \mu_n \sigma_n}{1+\mu_n^2 H^2} - \frac{A_p \mu_p \sigma_p}{1+\mu_p^2 H^2} \right) - \frac{\mu_n \zeta_n \sigma_n}{1+\mu_n^2 H^2} - \frac{\mu_p \zeta_p \sigma_p}{1+\mu_p^2 H^2}, \quad (10)$$

where the electron conductivity mobility $e\mu_n = \varepsilon\tau_n/m_n$, and the hole mobility $e\mu_p = \varepsilon\tau_p/m_p$.

From the equations (3), (4), (5), and (6), which are four independent relations involving the four unknowns σ_n , σ_p , μ_n , and μ_p , we find that

$$\mu_n = \frac{\Phi_{ET}(0) - \Phi_{ET}(\Omega)(1 + \mu_p^2 H^2)}{\Phi_{EL}(0) - \Phi_{EL}(\Omega)(1 + \mu_p^2 H^2)}, \quad \dots \quad (11)$$

$$\mu_p = -\frac{\Phi_{ET}(0) - \Phi_{ET}(\Omega)(1 + \mu_n^2 H^2)}{\Phi_{EL}(0) - \Phi_{EL}(\Omega)(1 + \mu_n^2 H^2)}. \quad \dots \quad (12)$$

Equations (11) and (12) may be solved by successive approximations.

The partial conductivities σ_n , σ_p are then obtainable from (3) and (5). The number (n) of conduction electrons and the number (p) of conduction holes per unit volume may also be deduced, since (see Wilson 1953, p. 198)

$$n = \sigma_n / \varepsilon c \mu_n, \quad p = \sigma_p / \varepsilon c \mu_p. \quad \dots \quad (13)$$

The substitution of μ_n , μ_p , σ_n , and σ_p in (7), (8), (9), and (10) yields four linear equations which may be solved for the four unknowns A_n , A_p , ζ_n , and ζ_p .

The parameters τ_n , τ_p , m_n , and m_p are then obtainable, since from (3), (4), and (5) we have

$$\tau_n = \frac{(c \mu_n \sigma_n^2)^{\frac{1}{2}}}{\varepsilon^{\frac{1}{2}} \zeta_n} \left(\frac{3 \hbar^3}{16 \sqrt{2} \pi \varepsilon^2} \right)^{\frac{2}{3}},$$

$$m_n = \varepsilon \tau_n / \mu_n c,$$

with similar expressions for τ_p and m_p .

Finally, the difference in energy (ΔE) between the base of the normal band and the top of the inverted band may be deduced from the relation $\Delta E = A_n - A_p$.

(ii) *Semiconductors*.—In many semiconductors the electrons and holes form non-degenerate systems ($\zeta \ll 0$). In this case the most convenient parameters are:

- (a) The electron and hole densities n and p .
- (b) The conductivity mobilities μ_n and μ_p .

As we have assumed (Section II) that both electrons and holes have Fermi-Dirac distributions, n and p are given by (Wilson 1953, p. 15)

$$\left. \begin{aligned} n &= -\frac{16 \sqrt{2} \pi m_n^{3/2}}{3 \hbar^3} \int_0^\infty E_n^{3/2} \frac{\partial f_{0n}}{\partial E_n} dE_n, \\ p &= -\frac{16 \sqrt{2} \pi m_p^{3/2}}{3 \hbar^3} \int_0^\infty E_p^{3/2} \frac{\partial f_{0p}}{\partial E_p} dE_p. \end{aligned} \right\} \quad \dots \quad (14)$$

In terms of the integral $I_{st}(x)$ (Section III (a)), equations (14) may be written

$$\left. \begin{aligned} n &= m_n \varepsilon^{-1} I_{01n}(0), \\ p &= m_p \varepsilon^{-1} I_{01p}(0). \end{aligned} \right\} \quad \dots \quad (15)$$

The mobilities are defined by (Shockley 1950, p. 16)

$$\sigma^{K0} = \varepsilon (n \mu_n + p \mu_p).$$

From (3.3) and (3.11) we have

$$\sigma^{K0} = \Phi_{EL}(0) = \varepsilon[I_{11n}(0) + I_{11p}(0)], \quad \dots\dots\dots (16)$$

and hence

$$\left. \begin{aligned} \mu_n &= I_{11n}(0)/n = \varepsilon I_{11n}(0)/m_n I_{01n}(0), \\ \mu_p &= I_{11p}(0)/p = \varepsilon I_{11p}(0)/m_p I_{01p}(0). \end{aligned} \right\} \quad \dots\dots\dots (17)$$

The above four parameters may be determined from Φ_{EL} and Φ_{ET} as we now proceed to show. On combining (3.3), (3.4), and (16) we obtain for these functions

$$\Phi_{EL}(\Omega) = \varepsilon \{ n \mu_n [L_{11n}(\Omega_n)/L_{11n}(0)] + p \mu_p [L_{11p}(\Omega_p)/L_{11p}(0)] \}, \quad \dots\dots (18)$$

$$\Phi_{ET}(\Omega) = \frac{\varepsilon}{c} \left[n \mu_n^2 \frac{L_{01n}(0)L_{21n}(\Omega_n)}{L_{11n}^2(0)} - p \mu_p^2 \frac{L_{01p}(0)L_{21p}(\Omega_p)}{L_{11p}^2(0)} \right]. \quad \dots\dots (19)$$

It is then necessary to evaluate the integral

$$L_{st}(x) = - \int_0^\infty \frac{\tau^s E^{t+\frac{1}{2}}}{1+x^2\tau^2} \frac{\partial f_0}{\partial E} dE.$$

For non-degenerate systems, $f_0 = \exp [(\zeta - E)/kT]$, and therefore (unlike case (i)) the energy dependence of τ must be taken into account. This varies with the scattering mechanism. Here, we consider in detail only the special case of scattering by the thermal vibrations of the lattice ($\tau \propto E^{-\frac{1}{2}}$). For ionic-impurity scattering [$\tau = g(E)E^{3/2}$] an exactly similar method is applicable, provided the slowly varying function $g(E)$ is replaced by a constant $g(E_m)$ —for a further discussion of this point see Dingle (1955), Beer, Armstrong, and Greenberg (1957)—and only the numerical constants differ in the final results. The third important case of mixed scattering is considerably more complicated since one further parameter (the ratio of thermal to ionic scattering) must also be evaluated.

For thermal scattering we have $\tau \propto E^{-\frac{1}{2}}$ (Wilson 1953, p. 265) and therefore L_{st} may be written

$$L_{st}(x) = (kT)^{t+\frac{1}{2}} a^s e^{\zeta/kT} \int_0^\infty \frac{y^{t+3/2-s/2}}{y+x^2a^2} e^{-y} dy \quad \dots\dots\dots (20)$$

$$= (kT)^{t+\frac{1}{2}} a^s e^{\zeta/kT} \left(t + \frac{3}{2} - \frac{1}{2}s \right)! A_{t+3/2-s/2}(x^2a^2), \quad \dots\dots\dots (21)$$

where $y = E/kT$ and $\tau = ay^{-\frac{1}{2}}$. The integral

$$A_r(u) = \frac{1}{r!} \int_0^\infty \frac{y^r}{y+u} e^{-y} dy$$

has been tabulated by Dingle, Arndt, and Roy (1957a).

For the two special cases of $x=0$ and $x\tau \gg 1$, equation (21) reduces to

$$L_{st}(0) = (kT)^{t+\frac{1}{2}} a^s e^{\zeta/kT} \left(t + \frac{1}{2} - \frac{1}{2}s \right)! \quad \dots\dots\dots (22)$$

$$L_{st}(x) = (kT)^{t+\frac{1}{2}} a^{s-2} e^{\zeta/kT} x^{-2} \left(t + \frac{3}{2} - \frac{1}{2}s \right)! \quad \dots\dots\dots (23)$$

On combining (17), (18), (19), (22), and (23) we obtain

$$\Phi_{EL}(0) = \varepsilon(n\mu_n + p\mu_p), \quad \dots\dots\dots (24)$$

$$H^2\Phi_{EL}(\Omega) = \frac{32c^2\varepsilon}{9\pi} \left(\frac{n}{\mu_n} + \frac{p}{\mu_p} \right) \quad \text{for } \Omega^2\tau^2 \gg 1, \quad \dots\dots (25)$$

$$\Phi_{ET}(0) = (3\pi\varepsilon/8c)(n\mu_n^2 - p\mu_p^2), \quad \dots\dots\dots (26)$$

$$H^2\Phi_{ET}(\Omega) = \varepsilon c(n - p) \quad \text{for } \Omega^2\tau^2 \gg 1. \quad \dots\dots\dots (27)$$

These four algebraic equations can be solved for n , p , μ_n , and μ_p .

The values of the parameters A_n , A_p may then be deduced from Ψ_{EL} and Ψ_{ET} since equations (3.7) and (3.8) reduce to

$$\begin{aligned} \Psi_{EL}(0) &= -A_n n \mu_n - A_p p \mu_p - I_{12n}(0) + I_{12p}(0) \\ &= -(A_n + 2kT)n\mu_n - (A_p - 2kT)p\mu_p, \quad \dots\dots\dots (28) \end{aligned}$$

$$\begin{aligned} \Psi_{ET}(0) &= -(3\pi/8c)(A_n n \mu_n^2 - A_p p \mu_p^2) - \varepsilon c^{-1}[m_n^{-1}I_{22n}(0) + m_p^{-1}I_{22p}(0)] \\ &= -(3\pi/8c) \left[\left(A_n + \frac{3}{2}kT \right) n \mu_n^2 - \left(A_p - \frac{3}{2}kT \right) p \mu_p^2 \right]. \quad \dots\dots\dots (29) \end{aligned}$$

Finally, the energy gap (ΔE) between the bands may be deduced from the relation $\Delta E = A_n - A_p$.

ON THE CYLINDRICAL PROBE METHOD OF MEASURING THERMAL CONDUCTIVITY WITH SPECIAL REFERENCE TO SOILS

I. EXTENSION OF THEORY AND DISCUSSION OF PROBE CHARACTERISTICS

By D. A. DE VRIES* and A. J. PECK*

[*Manuscript received December 3, 1957*]

Summary

The theory of cylindrical probes for measuring thermal conductivity is extended to the case of a probe of finite conductivity containing a line source at its centre. This provides a more realistic approximation to most actual probes than the theory for a probe of infinite conductivity developed by other authors. New experimental results are presented which are in complete agreement with theory. It is shown how an estimate can be obtained of the magnitude of a possible thermal contact resistance between the probe and the medium and how its influence on the measured conductivity can be assessed.

Conditions under which the theory of the infinite line source can be applied with a sufficient degree of accuracy are treated. The properties of various probes described in the literature are reviewed in this respect. The importance of measuring both the heating and the cooling branch of the temperature against time curve is emphasized. Difficulties in measuring thermal diffusivity with the probe are briefly discussed.

I. INTRODUCTION

Measurement of the thermal conductivity of soils and thermal insulating materials has received much attention over the past 10 years. One can distinguish between two groups of investigators interested in the subject who apparently have not always been aware of each other's activities in this field (see for instance a recent discussion in *Nature*: de Vries (1956), Webb (1956, 1957), Makowski and Mochlinski (1957)). The first group is that of engineers, who are interested in thermal properties of the soil in connexion with heat transfer from buried cables and coils of heat pumps, road construction, etc. The second group consists of soil scientists, hydrologists, meteorologists, oceanographers, and agronomists, who are concerned with such problems as the energy balance of the Earth's surface, the temperature regime of the upper soil and lower air layers, and the measurement of soil moisture content by thermal methods.

Steady state methods are not very suitable for use with soils. They cannot be applied *in situ* and in addition the soil water will redistribute itself under the influence of a temperature gradient. Various non-stationary methods have been proposed in which the temperature rise is measured of heated test bodies of various shapes which are inserted in the soil (Chudnowskii 1946, 1954; Skeib 1950; Misener 1952).

* Division of Plant Industry, C.S.I.R.O., Deniliquin, N.S.W.

Most actual measurements have been made with needle-shaped cylindrical test bodies (which we shall call cylindrical probes or, briefly, probes). These probes contain as a heat source a thin metal wire which is heated electrically; the temperature rise is measured by means of a thermocouple with its "warm" junction inside the probe near its centre. They can be installed *in situ* fairly easily without appreciable disturbance of the soil near the measuring thermojunction. When properly constructed and dimensioned the theoretical interpretation of the results is simple. In its simplest form the probe consists of a single heating wire, the temperature of which is measured by means of a thermocouple or by a resistance method.

The cylindrical probe method was first suggested by Schleiermacher (1888) and independently by Stålhane and Pyk (1931). The method was developed and used for measuring the thermal conductivity of liquids by Weishaupt (1940) and by van der Held and van Drunen (1949). The work of the latter authors served as the basis for many further developments. More recent work on the thermal conductivity of liquids was published by van der Held, Hardebol, and Kalshoven (1953), Gillam and Lamm (1955), Gillam *et al.* (1955), Hill (1957).

The method was first applied to soils by Hooper and Lepper (1950) and independently by Skeib (1950). Later work on soils was published by Hooper (1952), Mason and Kurtz (1952), de Vries (1952*a*, 1952*b*, 1953), van Duin and de Vries (1954), de Vries and de Wit (1954), Buettner (1955*a*), Makowski and Mochlinski (1956). Closely related to the work on soils is that of Bullard, Maxwell, and Revelle (1956) on sediments deposited on the ocean floor.

Cylindrical probes were used to measure the conductivity of thermal insulating materials by Hooper and Lepper (1950), d'Eustachio and Schreiner (1952), Mann and Forsyth (1956).

Measurements on rocks were published by Beck, Jaeger, and Newstead (1956).

The interpretation of the measurements is in most cases based on the theory of the infinite line source, supplemented by theoretical estimates or calculations of the influence of finite dimensions. A first attempt by van der Held and van Drunen (1949) to calculate the influence of the finite (i.e. non-zero) probe radius was unsatisfactory. This was recognized subsequently by van der Held, Hardebol, and Kalshoven (1953) and independently by Blackwell (1954). These authors treated the theory for a probe of finite radius and infinite thermal conductivity. The theory for this case was elaborated by Jaeger (1956). Blackwell (1954) also discussed the case of a hollow probe of finite conductivity with heat supplied at its surface. In a later paper Blackwell (1956) treated the influence of the finite probe length.

In the present paper the theory is extended to the case of a homogeneous probe of finite conductivity with a line source of heat at its centre. This gives a more realistic approximation to most actual probes than that of preceding theories. A detailed interpretation of experimental results obtained on dry sand with probes of various construction is given on the basis of this theory.

Following the discussion of probe theory and characteristics in this paper we shall treat the complications due to moisture effects in experiments on moist soils in a second paper.

II. NOTATION AND UNITS

- a , Thermal diffusivity of medium ($\text{cm}^2 \text{sec}^{-1}$),
- a_i , thermal diffusivity of probe ($\text{cm}^2 \text{sec}^{-1}$),
- c , specific heat ($\text{cal g}^{-1} \text{°C}^{-1}$),
- $C = \rho c$, volumetric heat capacity of medium ($\text{cal cm}^{-3} \text{°C}^{-1}$),
- C_i , volumetric heat capacity of probe ($\text{cal cm}^{-3} \text{°C}^{-1}$),
- H , heat transfer coefficient ($\text{cal cm}^{-2} \text{sec}^{-1} \text{°C}^{-1}$),
- I_n , modified Bessel function of first kind and order $n (=0,1)$,
- K_n , modified Bessel function of second kind and order $n (=0,1)$,
- l , half length of probe (cm),
- p , variable in Laplace transformation (sec^{-1}),
- $q = (p/a)^{1/2}$ (cm^{-1}),
- $q_i = (p/a_i)^{1/2}$ (cm^{-1}),
- Q , heat production per unit length of probe ($\text{cal cm}^{-1} \text{sec}^{-1}$),
- r , radial distance from axis of probe (cm),
- R , radius of probe (cm),
- R_1 , outer radius of soil sample (cm),
- t , time (sec),
- t_1 , time at end of heating (sec),
- T , temperature (°C),
- T_0 , initial temperature (°C),
- T_i , probe temperature (°C),
- $\alpha = \lambda/\lambda_i$,
- $\beta = C_i/C$,
- $\gamma = 0.5772$, Euler's constant,
- δ , thickness of air gap (cm),
- $\eta = \lambda/RH$, dimensionless contact resistance,
- λ , thermal conductivity of medium ($\text{cal cm}^{-1} \text{sec}^{-1} \text{°C}^{-1}$),
- λ_{air} , thermal conductivity of air ($\text{cal cm}^{-1} \text{sec}^{-1} \text{°C}^{-1}$),
- λ_i , thermal conductivity of probe ($\text{cal cm}^{-1} \text{sec}^{-1} \text{°C}^{-1}$),
- ρ , density (g cm^{-3}),
- $\tau = at/R^2$,
- $\tau_1 = at_1/R^2$,
- $\tau_i = at/l^2$.

III. PROBE THEORY

(a) *The Infinite Line Source*

In its simplest form the theory of the cylindrical probe is based on that of the infinite line source embedded in an infinite, homogeneous, isotropic medium. The Fourier equation of heat conduction can then be written

$$\frac{\partial T}{\partial t} = \alpha \left(\frac{\partial^2 T}{\partial r^2} + \frac{1}{r} \frac{\partial T}{\partial r} \right), \quad \dots \dots \dots (1)$$

The initial and boundary conditions are :

$$T=T_0, \text{ for } t=0, \text{ and for } r=\infty \text{ (} t \text{ finite),} \quad \dots\dots (2)$$

$$Q=-\lim_{r \rightarrow 0} 2\pi\lambda r \partial T / \partial r, \text{ for } t \geq 0. \quad \dots\dots\dots (3)$$

The solution to this problem is (Carslaw and Jaeger 1948)

$$T-T_0=(Q/4\pi\lambda)[-Ei(-r^2/4at)]. \quad \dots\dots\dots (4)$$

The exponential integral can be expanded as

$$-Ei(-x)=-\gamma-\ln x+x-\frac{1}{2}x^2+0(x^3). \quad \dots\dots\dots (5)$$

Hence for $4at/r^2 \gg 1$ we have to a good degree of approximation

$$T-T_0=(Q/4\pi\lambda)(-\gamma+\ln t+\ln 4a/r^2). \quad \dots\dots\dots (6)$$

When the source discontinues to operate at time t_1 we must replace condition (3) by

$$Q=-\lim_{r \rightarrow 0} 2\pi\lambda r \partial T / \partial r, \text{ for } 0 \leq t \leq t_1; \quad Q=0, \text{ for } t > t_1. \quad \dots (3a)$$

The solution now becomes

$$T-T_0=(Q/4\pi\lambda)[-Ei(-r^2/4at)+Ei\{-r^2/4a(t-t_1)\}], \text{ for } t \geq t_1, \dots (4a)$$

or, for $4a(t-t_1)/r^2 \gg 1$,

$$T-T_0=(Q/4\pi\lambda) \ln t/(t-t_1). \quad \dots\dots\dots (6a)$$

As we shall see below, a line source can be realized to a good degree of approximation by a thin metal wire (e.g. diameter 0.01 cm) which is heated by an electric current during the time interval 0 to t_1 . In that case λ can be found from an experiment by plotting $T-T_0$ against $\ln t$ for $t \leq t_1$, and also by plotting $(Q/4\pi\lambda)(-\gamma+\ln 4at/r^2)-(T-T_0)$ against $\ln(t-t_1)$. Values of the first term in this expression for $t > t_1$ are found by extrapolation of the line observed for $t < t_1$. The derivation of λ from both the heating and the cooling branches of the temperature-time curve provides a useful check on the procedure. This is of special importance in the case of measurements on moist porous media as will be discussed in Part II (de Vries and Peck 1958, in press).

(b) *The Probe of Finite Thickness containing a Line Source*

We shall now extend the theory to that of a homogeneous, isotropic, cylindrical probe of infinite length and radius R , which has an infinite line source at its centre.

The heat conduction problem must then be solved for a composite medium. The basic differential equation is again equation (1), this time with the thermal diffusivity a for $r > R$ and a_i for $r < R$. Instead of the boundary condition (3) or (3a) we now have a similar condition with λ_i substituted for λ . Finally, there are additional boundary conditions at the probe-medium interface. Assuming a contact resistance at this boundary with heat transfer coefficient H , we have

$$H[T_i(R,t)-T(R,t)]=-\lambda_i \partial T_i / \partial r = -\lambda \partial T / \partial r, \text{ for } r=R. \quad \dots (7)$$

Here $T_i(R, t)$ and $T(R, t)$ are respectively the temperatures of the probe and the medium at the interface.

The solution to this problem can be found by the method of the Laplace transformation. Denoting transformed temperatures by a bar the transformed equations are in the usual notation :

$$\frac{d^2 \bar{T}_i}{dr^2} + \frac{1}{r} \frac{d \bar{T}_i}{dr} - \frac{p \bar{T}_i}{a_i} = 0, \quad \text{for } 0 \leq r < R, \quad \dots \dots \dots (8)$$

$$\frac{d^2 \bar{T}}{dr^2} + \frac{1}{r} \frac{d \bar{T}}{dr} - \frac{p \bar{T}}{a} = 0, \quad \text{for } r > R, \quad \dots \dots \dots (9)$$

$$Q/p = - \lim_{r \rightarrow 0} 2\pi \lambda_i r d\bar{T}_i/dr, \quad \dots \dots \dots (10)$$

$$H(\bar{T}_i - \bar{T}) = -\lambda_i d\bar{T}_i/dr = -\lambda d\bar{T}/dr, \quad \text{for } r = R. \quad \dots (11)$$

The solution is :

$$\bar{T}_i - T_0/p = A I_0(q_i r) + Q K_0(q_i r)/2\pi \lambda_i p, \quad \dots \dots \dots (12)$$

$$\bar{T} - T_0/p = Q a_i^{1/2} K_0(qr)/2\pi \Delta p^{3/2} R, \quad \dots \dots \dots (13)$$

with

$$A = (Q/2\pi \Delta p) [a_i^{1/2} K_0(qR) K_1(q_i R) + a_i^{1/2} \eta q R K_1(qR) K_1(q_i R) - \lambda a_i^{1/2} \lambda_i^{-1} K_0(q_i R) K_1(qR)], \quad \dots \dots \dots (14)$$

$$\Delta = \lambda a_i^{1/2} I_0(q_i R) K_1(qR) + \lambda_i a_i^{1/2} I_1(q_i R) [K_0(qR) + \eta q R K_1(qR)], \quad \dots \dots \dots (15)$$

where $\eta = \lambda/RH$.

The temperatures T_i and T can now be found by applying an inverse Laplace transformation.* However, numerical calculation of T_i and T from the resulting analytical expressions is very laborious due to the oscillatory character of the Bessel functions occurring in the solutions.

Here we are principally interested in the behaviour of the solution for large values of time ($4at/R^2 \gg 1$). This can be found along the lines set out by Blackwell (1954). The expressions for \bar{T}_i and \bar{T} are expanded in ascending powers of p and the resultant series is integrated term by term along a contour in the p -plane which is cut along the negative real axis. This contour follows the lower negative axis ($xe^{-\pi i}$) from $-\infty$ to the origin, circles the origin counter-clockwise, and returns to $-\infty$ along the upper negative axis ($xe^{\pi i}$). Using the contour integrals listed by Blackwell and

$$\frac{1}{2\pi i} \int p \ln(cp) e^{tp} dp = \frac{1}{t^2}, \quad \dots \dots \dots (16)$$

$$\frac{1}{2\pi i} \int p \ln^2(cp) e^{tp} dp = -\frac{2}{t^2} \left(\ln \frac{t}{c} + \gamma - 1 \right), \quad \dots \dots \dots (17)$$

$$\frac{1}{2\pi i} \int p \ln^3(cp) e^{tp} dp = \frac{1}{t^2} \left[3 \left(\ln \frac{t}{c} + \gamma \right)^2 - 6 \left(\ln \frac{t}{c} + \gamma \right) - \frac{1}{2} \pi^2 \right] \quad \dots (18)$$

* D. A. de Vries, unpublished result (for $\eta=0$ only).

(where the integration is along the contour and c is a real constant), we obtain :

$$T_i(r, t) - T_0 = (Q/4\pi\lambda)G_i(\tau, \eta, \alpha, \beta, r/R), \quad \text{for } r \leq R, \quad \dots \quad (19)$$

$$T(R, t) - T_0 = (Q/4\pi\lambda)G(\tau, \eta, \alpha, \beta), \quad \dots \quad (20)$$

with $\alpha = \lambda/\lambda_i$, $\beta = C_i/C$, and $\tau = 4at/R^2$. The functions G_i and G are in dimensionless form :

$$G_i = \ln \tau - \gamma + 2\eta - 2\alpha \ln(r/R) + c_{i,-1}\tau^{-1} + 0(\tau^{-2} \ln^2 \tau), \quad \dots \quad (21)$$

with

$$c_{i,-1} = 2(1-\beta)(\ln \tau - \gamma) + 2 - 4\eta\beta - 2\alpha\beta + \alpha\beta r^2/R^2, \quad \dots \quad (22)$$

and

$$G = \ln \tau - \gamma + c_{-1}\tau^{-1} + c_{-2}\tau^{-2} + 0(\tau^{-3} \ln^3 \tau), \quad \dots \quad (23)$$

with

$$c_{-1} = 2(1-\beta)(\ln \tau - \gamma) + 2 - 2\eta\beta - \alpha\beta, \quad \dots \quad (24)$$

$$\begin{aligned} c_{-2} = & -3(1-\beta)^2(\ln \tau - \gamma)^2 + (-1 - 4\beta + 6\beta^2 + 2\alpha\beta - 3\alpha\beta^2 + 4\eta\beta - 8\eta\beta^2)(\ln \tau - \gamma) \\ & + \frac{1}{2}\pi^2(1-\beta)^2 + \frac{3}{2} - 4\beta + 3\alpha\beta^2 - \frac{3}{4}\alpha^2\beta^2 + 8\eta\beta^2 - 3\eta\alpha\beta^2 - 4\eta^2\beta^2. \quad \dots \quad (25) \end{aligned}$$

It can be easily checked that G_i and G reduce to (5) for $\alpha = \beta = 1$ and $\eta = 0$. For a probe of infinite conductivity ($\alpha = 0$) the expression for G_i reduces to those given by Blackwell (1954) and Jaeger (1956) for large τ . Terms of the order τ^{-2} were not computed for G_i because of the unwieldiness of the resulting expressions. However, in the applications considered below it is safe to assume that the term with τ^{-2} in G_i is of the same order of magnitude as that in G . A further discussion of these equations is given in Section IV.

Equations (19) and (20) hold for the heating branch only. For the cooling branch ($t > t_1$) we have, analogous to equation (4a) :

$$T_i(r, t) - T_0 = (Q/4\pi\lambda)[G_i(\tau, \eta, \alpha, \beta, r/R) - G_i(\tau - \tau_1, \eta, \alpha, \beta, r/R)], \quad \dots \quad (26)$$

$$T(R, t) - T_0 = (Q/4\pi\lambda)[G(\tau, \eta, \alpha, \beta) - G(\tau - \tau_1, \eta, \alpha, \beta)]. \quad \dots \quad (27)$$

(c) Miscellaneous Factors

In the interpretation of a thermal conductivity experiment we must also consider systematic errors arising from various factors apart from the finite radius of the probe. These factors are : finite length of the probe, finite dimensions of the sample, the inertia of the temperature-measuring system, and factors connected with moisture movement and an uneven distribution of moisture in the sample. The moisture problems will be discussed in a separate paper ; a brief discussion of the other factors is given in this subsection.

The influence of the finite probe length can be made arbitrarily small, of course, by a proper choice of the probe dimensions. Blackwell (1956) has derived an upper limit for the relative error in the slope of the T_i against $\ln t$ curve due to axial heat flow on the assumptions that the probe is a good conductor

and that heat is generated uniformly throughout the probe. His equation for this upper limit reads in our notation

$$1 - \frac{\partial T_i(R, t) / \partial \ln t}{Q / 4\pi\lambda} = \pi^{-\frac{1}{2}} \left[\tau_l^{\frac{1}{2}} + \frac{2(1 - \alpha\beta)}{\alpha\tau_l^{\frac{1}{2}}\tau} (\ln \tau - \gamma + 2\eta) \right] \exp(-\tau_l^{-1}), \dots (28)$$

with $\tau_l = 4at/l^2$, where l is half the probe length. Once the probe diameter and the heating time are fixed a proper value of l can be derived from equation (28). As we shall see in the following section, this does not lead to excessive probe lengths.

The influence of the finite dimensions of the soil sample can also be made sufficiently small without difficulty by a proper choice of the dimensions of the sample container. We propose the following simple criterion for this purpose: the amount of heat passing through the walls of the container must be small in comparison with the heat input at the source during the time of heating. Since we are only concerned with its order of magnitude a sufficiently accurate estimate of this amount of heat can be obtained by application of the solution for an infinite line source. For a long cylindrical sample container with radius R_1 this leads to the inequality

$$\exp(-R_1^2/4at_1) \ll 1. \dots\dots\dots (29)$$

The temperature is usually measured near the centre of the source by a thermocouple in connexion with a galvanometer. The inertia of this system will be negligible for the type of temperature variation encountered here when the time at which the first reading is taken is large in comparison with the period of the galvanometer. In our experiments the former was not less than 10 sec and the latter was 0.2 sec. In this case the error in the measured deflection is less than 1 per cent.

IV. NUMERICAL AND EXPERIMENTAL DATA

In this section we shall apply the preceding theory in discussing the various current methods for measuring the thermal conductivity of soils by means of heated cylindrical test bodies. The methods used previously by one of us (de Vries 1952*a*, 1952*b*) are treated in detail and new experimental results are presented to illustrate the argument. A brief discussion of probes described by other authors is given in subsection (*d*).

Thermal properties of various materials used in the construction of probes are listed in the upper part of Table 1, whilst those of soil materials and soils are given in the lower part of this table. The former were taken from handbooks of physical constants, the latter were derived from various sources (see de Vries (1952*b*) and de Vries and de Wit (1954)).

(*a*) Single Wire

The simplest form of a linear heat source is that of a straight metal wire heated by an electric current. We have used constantan and manganin wires with diameters of 0.01 to 0.02 cm for this purpose. The temperature was measured by means of a thermojunction as close as possible to the centre

of the wire. The thermocouple was made of copper and constantan wires with a diameter of 0.01 cm. All wires were enamelled to secure good electrical insulation between the heating circuit and the temperature-measuring circuit. This method was employed previously by one of us (de Vries 1952*b*) for laboratory determinations of the thermal conductivity of a sand at various temperatures and moisture contents.

TABLE 1
THERMAL PROPERTIES OF VARIOUS MATERIALS AT 20 °C

Material	Density, ρ (g cm ⁻³)	Specific Heat, c (cal g ⁻¹ °C ⁻¹)	Vol. Heat Capacity, C (cal cm ⁻³ °C ⁻¹)	Thermal Conductivity, λ (cal cm ⁻¹ sec ⁻¹ °C ⁻¹)	Thermal Diffusivity, a (cm ² sec ⁻¹)
Copper ..	8.89	0.092	0.82	0.92	1.12
Manganin	8.50	0.097	0.82	0.15	0.18
Constantan	8.88	0.099	0.88	0.054	0.061
Monel ..	8.90	0.098	0.87	0.052	0.060
Glass* ..	2.6	0.2	0.5	2.6×10^{-3}	5×10^{-3}
Paraffin ..	0.89	0.69	0.61	0.6×10^{-3}	1.0×10^{-3}
Air	0.0012	0.24	0.00029	0.062×10^{-3}	0.21
Quartz ..	2.65	0.175	0.46	0.020	0.043
Many soil minerals*	2.65	0.175	0.46	0.007	0.015
Soil organic matter* ..	1.3	0.46	0.60	0.6×10^{-3}	1.0×10^{-3}
Soil, mineral, dry* ..	1.50	—	0.26	0.5×10^{-3}	1.9×10^{-3}
Soil, mineral, saturated*	1.93	—	0.69	5×10^{-3}	7×10^{-3}
Soil, organic, dry* ..	0.13	—	0.060	0.08×10^{-3}	1.2×10^{-3}
Soil, organic, saturated*	1.03	—	0.96	1.2×10^{-3}	1.3×10^{-3}

* Approximate average values.

For a heating wire of sufficient length the temperature will lie between that given by equations (19) and (20) for the heating branch and equations (26) and (27) for the cooling branch, where in (19) and (26) we must take $r=R$. Strictly, we should apply similar equations derived for heat production throughout the wire instead of at its centre, but the differences are negligible here because the conductivity of the wire is large in comparison with that of the soil.

From the data given in Table 1 it can be easily checked that for R -values of 0.01 cm or less the terms of order τ^{-1} and higher negative order in equations (21) and (23) are negligible for $t > 10$ sec, unless η is very large, say greater than 10. Such large values of η are unlikely when soil is packed around the wire, and a numerical example given below leads to a value of the order of unity.

During our experiments the heating time, t_1 , is usually 180 sec. It follows from equation (28) that the relative error due to axial flow is less than 1 per cent. for $l > 5$ cm, the actual l -value being usually 7.5 cm. Finally R_1 in (29) was 5 cm or more, which, with $a \leq 10^{-2}$ cm² sec⁻¹ and $t_1 = 180$ sec, leads to $\exp(-R_1^2/4at_1) \leq 0.03$.

Results of an experiment conducted with a dry coarse quartz sand (particle sizes ranging from 0.060 to 0.085 cm) and with a manganin heating wire of 0.01 cm diameter are presented in Figure 1 (line *a*). The density of the sand was 1.54 g cm⁻³, its temperature 20 °C. It will be noted that deviations from

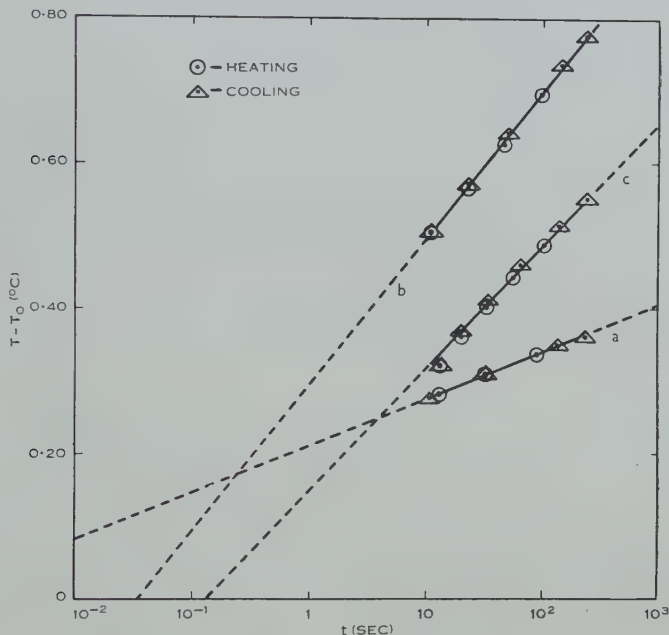


Fig. 1.—Temperature rise, $T - T_0$, against time, t , for thermal conductivity experiments. For the cooling branch time is counted from end of heating (180 sec) onwards and points are corrected for the influence of previous heating (see Section III (a)). Line *a*, single heating wire, heat input $Q = 2.3 \times 10^{-4}$ cal cm⁻¹ sec⁻¹; line *b*, heating wire in glass capillary, $Q = 8.0 \times 10^{-4}$ cal cm⁻¹ sec⁻¹; line *c*, probe, $Q = 6.5 \times 10^{-4}$ cal cm⁻¹ sec⁻¹.

the logarithmic relation are negligible. The value of λ following from this experiment was 0.67×10^{-3} cal cm⁻¹ sec⁻¹ °C⁻¹. Each of the three examples given under (a), (b), and (c) in this section are typical of several runs taken on a sample which gave the same results within the experimental accuracy (± 5 per cent. for λ).

From equation (21) with $r = R$ it follows that the intercept of an extrapolated line on the horizontal axis is

$$t = (R^2/4a)e^{\gamma-2\eta}, \quad \dots \dots \dots (30)$$

whilst according to equation (24) this intercept becomes

$$t = (R^2/4a)e^{\gamma}. \quad \dots \dots \dots (31)$$

The observed intercept will lie somewhere between these two theoretical values, probably close to the former. When $\eta=0$ the two intercepts coincide. In our example we find, with $\eta=0$, $C=0.27 \text{ cal cm}^{-3} \text{ }^{\circ}\text{C}^{-1}$, and $t=5.15 \times 10^{-4} \text{ sec}$, an R -value of 0.0017 cm . This is less than the actual value of 0.005 cm . Substituting the latter in (30) we obtain $\eta=1.08$.

A contact resistance will be due to an air gap between the heating wire and the medium. For an annular gap of thickness δ we have

$$\eta = (\lambda/\lambda_{\text{air}}) \ln (R+\delta)/R. \quad \dots\dots\dots (32)$$

In the present example this leads to a δ -value of $5 \times 10^{-4} \text{ cm}$, which is small in comparison with the soil particle size and with the diameter of the wire.

Apparently large positive or even negative values of η have been observed when the galvanometer received a leakage current from the heating circuit. Such a leakage is easily detected by reversing the direction of the heating current.

The methods used in the experiments given in subsections (a), (b), and (c) were previously described in detail by one of us (de Vries 1952a, 1952b). Here we shall by way of example give detailed information on the experiment with the single heating wire.

The resistance per unit length of the wire was determined at $0.610 \pm 0.005 \text{ } \Omega \text{ cm}^{-1}$, the heating current was measured at $0.0400 \pm 0.0005 \text{ A}$, which leads to a Q -value of $(2.33 \pm 0.08) \times 10^{-4} \text{ cal cm}^{-1} \text{ sec}^{-1}$. The time of heating was 180 sec.

The temperature rise was measured by means of a copper-constantan thermocouple connected to a micro-Moll galvanometer. The sensitivity of this system was $0.282 \pm 0.003 \text{ }^{\circ}\text{C}$ per cm deflection on the galvanometer scale. The time was measured with a stopwatch when a line-shaped light mark coincided with a millimetre division on the scale. The following results were obtained :

	Heating			Cooling			
Deflection (mm)	10	11	12	3	2	1	0.7
Time (sec) ..	13	33	90	191	213	315	420

The accuracy of the time readings was better than $\pm 0.5 \text{ sec}$. Towards the end of the cooling run the movement of the light mark became so slow that the uncertainty in the reading of the coincidence of the mark with a scale division was of the order of $\pm 5 \text{ sec}$. However, such an error represents only a small distance on the logarithmic time scale for the times concerned.

From a graph of $T - T_0$ against $\log t$ the slope of the resulting line is determined at $0.064 \pm 0.001 \text{ }^{\circ}\text{C}$, from which λ is found at $(0.67 \pm 0.03) \times 10^{-3} \text{ cal cm}^{-1} \text{ sec}^{-1} \text{ }^{\circ}\text{C}^{-1}$.

(b) Heating Wire in Glass Capillary

A thin heating wire is not always suitable for use in moist soils owing to the fact that relatively large values of the temperature gradient occur close to the wire. This can lead to a strong migration of moisture near the wire. We

therefore also use a heating wire fitted in a glass capillary with an external diameter of about 0.05 cm and internal diameter slightly greater than the wire. In that case the steepest temperature gradients are found in the glass instead of in the soil sample. The thermojunction is located outside the capillary as close as possible to its outer surface.

For R -values of about 0.025 cm and $t > 10$ sec the terms of order τ^{-1} and τ^{-2} are still negligible with mineral soils unless η is large. However, η will be smaller in this case than with the single wire and η will be negligibly small when the diameter of the probe is of the same magnitude as the sizes of the larger soil particles, assuming that the soil is well packed around the probe.

An exception must be made for organic soils of low density at low moisture contents (see Table 1), where the terms with τ^{-1} and τ^{-2} in (23) are not negligible in comparison with $\ln \tau - \gamma$, mainly due to the large value of β . In these loosely packed dry soils large values of η can also be expected.

Errors due to axial flow are again negligible for $l > 5$ cm. Equation (28) is not strictly applicable in this case because the conductivity of the probe is not necessarily large in comparison with that of the soil. However, this equation will still give a fair estimate of the error because in this case the dominant term is $\pi^{-1} \tau_l^{-1/2} \exp(-\tau_l^{-1})$. This term is due to deviations from radial flow in the (infinite) medium surrounding the probe and is independent of its diameter and thermal properties.

As an example results of an experiment with a heater of this type on the same coarse sand as mentioned above is given in Figure 1 (line *b*). The density of the dry sand was 1.58 g cm⁻³, its temperature 17 °C. The measured conductivity was 0.73 cal cm⁻¹ sec⁻¹ °C⁻¹. From the intercept on the horizontal axis (0.035 sec) and equation (31) we find $R = 0.014$ cm, which is smaller than the actual value, 0.030 cm. Substitution of the latter in (30) leads to $\eta = 0.38$ and from (32) to $\delta = 1.0 \times 10^{-3}$ cm.

(c) The Thermal Conductivity Probe

A probe for measuring thermal conductivity of soil *in situ* was developed by one of us (de Vries 1952*a*, 1952*b*). Its construction and dimensions are shown in Figure 2. An automatic recording apparatus to be used in connexion with the probes was described by van Duin and de Vries (1954).

The diameter of the probe is about 0.11 cm. Average values of λ_i and C_i will be used in finding α and β . From the composition of the probe its average volumetric heat capacity is found to be 0.63 cal cm⁻³ °C⁻¹. The average radial conductivity of the probe outside the heating wire is estimated at 1.0×10^{-3} cal cm⁻¹ sec⁻¹ °C⁻¹. It will be noted that the thermal resistance is mainly located in the paraffin. The distance of the thermojunction from the centre of the heat source is approximately 0.025 cm. Values of η will be negligibly small when there is a good contact between the soil and the probe, except for very coarse sand or soil of very low density.

It can be easily checked that for $t > 10$ sec terms of order τ^{-1} and of higher negative powers of τ in equations (21) and (22) are still small in comparison with

$\ln \tau - \gamma$ for mineral soils and for wet organic soils. A numerical example is given below. For very loose and dry organic soils these equations can no longer be applied, because the various terms become of the same order of magnitude. Apart from experiments with the latter soils (which are very rare), the theory of the line source can be applied to a probe of this construction to a sufficient approximation.

As an example we give the results obtained with a probe in the same coarse sand as mentioned before at a dry density of 1.54 g cm^{-3} and at 20°C (Fig. 1, line *c*). Slight deviations from the linear relationship occur with the first few points ($t=13$ and 20 sec), which are of the expected magnitude. On such

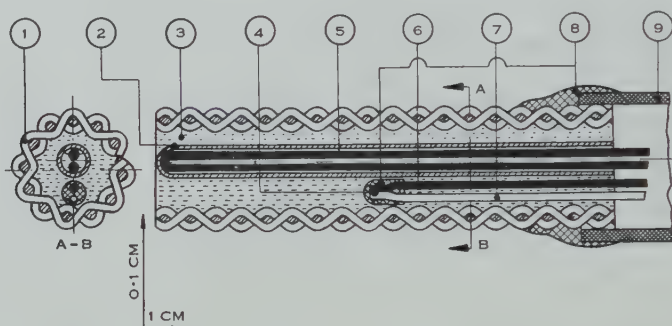


Fig. 2.—Radial and longitudinal cross sections of probe. 1, Monel gauze (filled with paraffin wax); 2, glass capillary; 3, paraffin wax; 4, thermojunction; 5, heating wire; 6, constantan wire; 7, copper wire; 8, insulating cover; 9, plastic socket.

occasions more weight is given to the points at higher t in drawing the line. In this case the value of λ was found to be $0.71 \times 10^{-3} \text{ cal cm}^{-1} \text{ sec}^{-1} ^\circ \text{C}^{-1}$. With $C=0.27 \text{ cal cm}^{-3} ^\circ \text{C}^{-1}$ this leads to $a=2.63 \times 10^{-3} \text{ cm}^2 \text{ sec}^{-1}$, $\alpha=0.71$, and $\beta=2.33$. At $t=13 \text{ sec}$ we find $\tau=45.1$.

From equation (21) with $\eta=0$ and the observed intercept on the horizontal axis ($t=0.13 \text{ sec}$) we find $r=0.021 \text{ cm}$, which is sufficiently close to the estimated value (0.025 cm) to support the assumption $\eta=0$. Using equation (21) with $t=13 \text{ sec}$ we have: $\ln \tau - \gamma - 2\alpha \ln r/R = 4.607$ and $c_{i-1}\tau^{-1} = -0.214$. The relative deviation from the straight line due to the term $c_{i-1}\tau^{-1}$ is therefore $-0.214/4.607 = -0.046$. In comparison, the observed value at $t=13$ is -0.045 (± 0.005 due to a possible random error in reading t); this includes, of course, the effect of terms with higher negative powers of τ . The values of the various terms on the right-hand side of equation (23) are in this case: $\ln \tau - \gamma = 3.234$, $c_{-1}\tau^{-1} = -0.172$, and $c_{-2}\tau^{-2} = 8.8 \times 10^{-5}$. With increasing τ , i.e. for longer times or for larger a (moister or denser soil), the deviations from the theory of the line source become even smaller.

According to equation (21) the intercept for $T=T_0$ is

$$t = (R^2/4a)(r/R)^{2\alpha} e^{\gamma - 2\eta}. \quad \dots \dots \dots (33)$$

From this equation a check can be obtained on the approximate value of η . This is of special importance in field experiments, where it would otherwise be impossible to know whether or not a sufficiently good thermal contact between the probe and the soil exists. A value of a based on an estimated value of C will suffice for this purpose. Such an estimate can be obtained from the dry density of the soil (measured on samples taken when placing the probes) and its approximate moisture content which follows from the measured conductivity.

When η is not negligible its influence on the value of λ can be assessed from equation (21). The term 2η has no influence on the slope of the $T-T_0$ against $\ln t$ curve. The contribution of η to the correction term of first order is

$$-\frac{4\eta\beta}{\tau} = -\frac{C_i R^2}{\lambda_{\text{air}} t} \ln \left(1 + \frac{\delta}{R}\right) \approx -\frac{C_i R \delta}{\lambda_{\text{air}} t}, \dots\dots\dots (34)$$

which is independent of the thermal properties of the medium. If, for instance, $\delta = 0.01$ cm we have $C_i R \delta / \lambda_{\text{air}} t = 1.12/t$ and the first order term in G_i due to η becomes -0.112 and -0.0063 after 10 and 180 sec respectively. The resulting relative error in λ is therefore $-0.106/\ln 18 = -0.037$. The contribution of the terms proportional to $\eta\tau^{-2}$ will be much smaller (see equation (25)).

The possible error in the λ -values following from the probe experiments was discussed by one of us (de Vries 1952b); it was found to be about ± 5 per cent. It was shown in the same paper that the λ -values for a dry quartz sand found from probe experiments were in close agreement with results on similar materials obtained with more conventional stationary methods by Smith and Byers (1938) and Kersten (1949). In addition, it was demonstrated that a theoretical calculation of the thermal conductivity from the composition of the sand led to values which were in good agreement with the experimental ones.

We therefore conclude that the probe method can be applied with confidence as an absolute method for measuring thermal conductivity. This conclusion receives further support from the work of others who compared probe results with those obtained by the guarded hot plate method (see d'Eustachio and Schreiner 1952; Mann and Forsyth 1956). Moreover, the material presented in this section shows that a detailed theoretical interpretation can be given of all aspects of the observed temperature curves.

(d) Other Probes, General Remarks

It will be clear from the foregoing examples that it is desirable to keep the probe diameter as small as possible to secure large values of τ and thereby small deviations from the simple logarithmic relationship between $T-T_0$ and t . These deviations decrease, of course, with increasing time, but large values of t are themselves undesirable. They necessitate the use of a long probe and large sample containers. In addition, with moist soils the amount of water that moves away from the probe increases with time.

Deviations from the simple theory due to the thickness of the probe can be further reduced by choosing α small and β close to unity. These requirements

cannot be met entirely in a single probe over the range of λ and C values found in soils. In some instances a hollow probe will be preferable to a solid one.

Various characteristics of a number of cylindrical probes designed for measuring the conductivity of soils and thermal insulating materials are listed in Table 2 in chronological order of publication. In most cases application of the simple theory will be permissible with these probes, except that of Buettner

TABLE 2

DIMENSIONS, THERMAL PROPERTIES, AND TIME INTERVALS FOR CYLINDRICAL PROBES

Authors	Diameter, $2R$ (cm)	Length, $2l$ (cm)	Radial Thermal Conductivity, λ_i (cal cm ⁻¹ sec ⁻¹ °C ⁻¹)	Vol. Heat Capacity, C_i (cal cm ⁻³ °C ⁻¹)	Time of First Reading t_0 (sec)	Time of Heating t_1 (sec)
Skeib (1950)	0.11	25	12×10^{-3}	0.38	50	100
Hooper and Lepper (1950) ..	0.48	46	0.08×10^{-3}	0.37	180	420
de Vries (1952 <i>a</i> , 1952 <i>b</i>) ..	0.11	13	1.0×10^{-3}	0.63	10	180
d'Eustachio and Schreiner (1952)† ..	0.076	10	0.6×10^{-3} *	0.60*	—	—
Mason and K u r t z (1952)† ..	0.63	60	—	—	60	1500
Buettner (1955 <i>a</i>) ..	0.07	2.5	1.0×10^{-3} *	0.60*	1	10
Bullard <i>et al.</i> (1956) ..	0.086	6.34	—	—	5	600
Mann and Forsyth (1956) ..	0.14	10	1.0×10^{-3} *	0.60*	20	120
Makowski and Mochlinski (1956) ..	0.48	46	0.08×10^{-3}	0.37	40	1200

* Value uncertain.

† Heater in the form of a coil.

(1955*a*), which is too short. Buettner (1955*b*) based his design on the theory for a probe of infinite conductivity ($\alpha=0$). However, his numerical values are in error (cf. Jaeger 1956) and, in addition, the approximation $\alpha=0$ is not permissible for the smaller τ -values in his experiments. This probably is the reason for his use of an empirical calibration constant to be obtained by calibrating the probe in materials with a C -value close to that of soil.

We wish to emphasize that the determination of more than two points on the $T-T_0$ against $\ln t$ curve is highly desirable. A procedure in which the slope of a line is determined from two points only (e.g. Hooper and Lepper 1950; Skeib 1950) provides no clues as to the occurrence of systematic or large random errors. On the other hand a reasonable safeguard is obtained by adopting the following procedure:

- (1) lines are plotted for both the heating and the cooling branches of the temperature curve, which should be found to coincide within the expected experimental error,
- (2) a check is made on the magnitude of a possible contact resistance by the method set out in subsection (c).

For certain applications it is impossible to achieve sufficiently large values of τ to apply the simple theory or large time approximations such as equations (21) and (23). This is the case, for instance, with measurements of rock conductivity in boreholes where the probe diameter must be rather large, and in measuring the conductivity of liquids where convection sets in after a certain time. It is then necessary to apply analytical solutions, which so far have only been derived and calculated numerically for infinite conductivity of the probe. In the results reported by van der Held, Hardebol, and Kalshoven (1953) the probe conductivity is not sufficiently great in comparison with that of the measured liquids for the approximation $\alpha=0$ to hold. This might account for the systematic error reported by these authors.

An interesting modification of the probe method for application to liquids was published recently by Hill (1957). In this method the thermocouple wires also serve to carry the heating current, the latter being applied during a fraction of a second only.

(c) *Measuring Thermal Diffusivity*

It has been suggested by some authors (Skeib 1950; Misener 1952) that the thermal diffusivity can also be obtained from a probe experiment, whereas others (Buettner 1955b; Beck, Jaeger, and Newstead 1956) have pointed to the difficulties of doing so.

From equation (21) it will be clear that a can only be found when λ_i , r , R , and η are known. R is easy to measure, whilst λ_i and r could be determined for each probe by calibration in two materials of known conductivity and diffusivity. The value of η depends on the contact between the probe and the soil, which during field experiments may change in the course of time. A determination of a only seems possible when it can reasonably be expected that η is negligibly small. But even then the accuracy in the value of a would be small, as it can be easily seen from the lines in Figure 1 that a small relative error in λ would cause a much larger relative error in the intercept on the horizontal axis and thereby in the measured diffusivity. With line c in Figure 1, for instance, a relative error of 3 per cent. in λ would result in a relative error of about 20 per cent. in a , assuming that the observed point at $t=100$ sec is correct.

V. CONCLUSIONS

It is shown theoretically that the thermal conductivity of soils and materials of similar thermal properties can be accurately measured by the cylindrical probe method. Experimental results in close agreement with the theory are presented.

The outer diameter of the probe should preferably be of the order of 0.1 cm or less, its length of the order of 10 cm. The volumetric heat capacity of the probe should not be large in comparison with that of the observed material, its thermal conductivity should preferably not be small in comparison with that of the material.

An estimate of the contact resistance coefficient can be obtained from the observed temperature rise and an estimate of the volumetric heat capacity of the material. The influence of the contact resistance on the value of the thermal conductivity can then be assessed from the theory; it is negligible when the soil is well packed around the probe.

The thermal diffusivity can be found when the contact resistance coefficient is negligibly small or accurately known, but no high degree of accuracy is to be expected.

VI. REFERENCES

- BECK, A., JAEGER, J. C., and NEWSTEAD, G. (1956).—*Aust. J. Phys.* **9**: 286–96.
 BLACKWELL, J. H. (1954).—*J. Appl. Phys.* **25**: 137–44.
 BLACKWELL, J. H. (1956).—*Canad. J. Phys.* **34**: 412–7.
 BUETTNER, K. (1955a).—*Trans. Amer. Geophys. Un.* **36**: 827–30.
 BUETTNER, K. (1955b).—*Trans. Amer. Geophys. Un.* **36**: 831–7.
 BULLARD, E. C., MAXWELL, A. E., and REVELLE, R. (1956).—*Advanc. Geophys.* **3**: 153–81.
 CARSLAW, H. S., and JAEGER, J. C. (1948).—“Conduction of Heat in Solids.” p. 221. (Clarendon Press: Oxford.)
 CHUDNOWSKII, A. F. (1946).—*Proc. Lenin. Acad. Agric. Sci. U.S.S.R.* **1946**: 38–43.
 CHUDNOWSKII, A. F. (1954).—*Pochvovedenie* **1954**: 64–71.
 VAN DUIN, R. H. A., and DE VRIES, D. A. (1954).—*Neth. J. Agric. Sci.* **2**: 168–75.
 D'EUSTACHIO, D., and SCHREINER, R. E. (1952).—*Heat. Pip. Air Condit.* **1952**: 113–7.
 GILLAM, D. G., and LAMM, O. (1955).—*Acta Chem. Scand.* **9**: 657–60.
 GILLAM, D. G., ROMBÉN, L., NISSEN, H. E., and LAMM, O. (1955).—*Acta Chem. Scand.* **9**: 641–56.
 VAN DER HELD, E. F. M., and VAN DRUNEN, F. G. (1949).—*Physica* **15**: 865–81.
 VAN DER HELD, E. F. M., HARDEBOL, J., and KALSHOVEN, J. (1953).—*Physica* **19**: 208–15.
 HILL, R. A. W. (1957).—*Proc. Roy. Soc. A* **239**: 476–86.
 HOOPER, F. C. (1952).—Highway Research Board, Special Rep. No. 2: 57–9.
 HOOPER, F. C., and LEPPER, F. R. (1950).—*Heat. Pip. Air Condit.* **1950**: 129–35.
 JAEGER, J. C. (1956).—*Aust. J. Phys.* **9**: 167–79.
 KERSTEN, M. S. (1949).—Bull. Univ. Minn. Engng. Exp. Sta. No. 28.
 MAKOWSKI, M. W., and MOCHLINSKI, K. (1956).—*Proc. Instn. Elect. Engrs.* **A 103**: 453–70.
 MAKOWSKI, M. W., and MOCHLINSKI, K. (1957).—*Nature* **179**: 778–9.
 MANN, G., and FORSYTH, F. G. (1956).—*Mod. Refrig.* **59**: 188–91.
 MASON, V. V., and KURTZ, M. (1952).—*Trans. Amer. Inst. Elect. Engrs.* **71**: 570–7.
 MISENER, A. D. (1952).—Highway Research Board, Special Rep. No. 2: 51–7.
 SCHLEIERMACHER, A. L. E. F. (1888).—*Wied. Ann. Phys.* **34**: 623.
 SKEIB, G. (1950).—*Z. Met.* **4**: 32–9.
 SMITH, W. O., and BYERS, H. G. (1938).—*Proc. Soil Sci. Soc. Amer.* **3**: 13–9.
 STÅLHANE, B., and PYK, S. (1931).—*Tekn. Tidskr.* **61**: 389–93.
 DE VRIES, D. A. (1952a).—*Soil Sci.* **73**: 83–9.

- DE VRIES, D. A. (1952b).—*Meded. Landbhogesch., Wageningen* **52**: 1-72.
- DE VRIES, D. A. (1953).—*Neth. J. Agric. Sci.* **1**: 115-21.
- DE VRIES, D. A. (1956).—*Nature* **178**: 1074.
- DE VRIES, D. A., and PECK, A. J. (1958).—On the cylindrical probe method of measuring thermal conductivity with special reference to soils. II. Analysis of moisture effects. *Aust. J. Phys.* **11** (3). (In press.)
- DE VRIES, D. A., and DE WIT, C. T. (1954).—*Met. Rdsch.* **7**: 41-5.
- WEBB, J. (1956).—*Nature* **178**: 1074.
- WEBB, J. (1957).—*Nature* **179**: 779.
- WEISHAUP, J. (1940).—*ForschArb. IngWes.* **11**: 20-35.

SHORT COMMUNICATIONS

LUNAR TIDES IN E_{2s} AT BRISBANE*

By A. D. GAZZARD†

From an analysis of the Brisbane $h'E_s$ data for 1953, Thomas and Svenson‡ have deduced the existence of a lunar semi-diurnal tide, with amplitude 0.69 km, and epoch of maximum at 6.7 hr local time. The present author has made a similar analysis of the Brisbane $h'E_{2s}$ data for the two years, 1952–1953, using nearly 3000 hourly readings.

The result is shown by the harmonic dial (Fig. 1). The harmonic coefficients for the 22 calendar months are plotted (Δ), as well as the computed mean, shown (\odot).

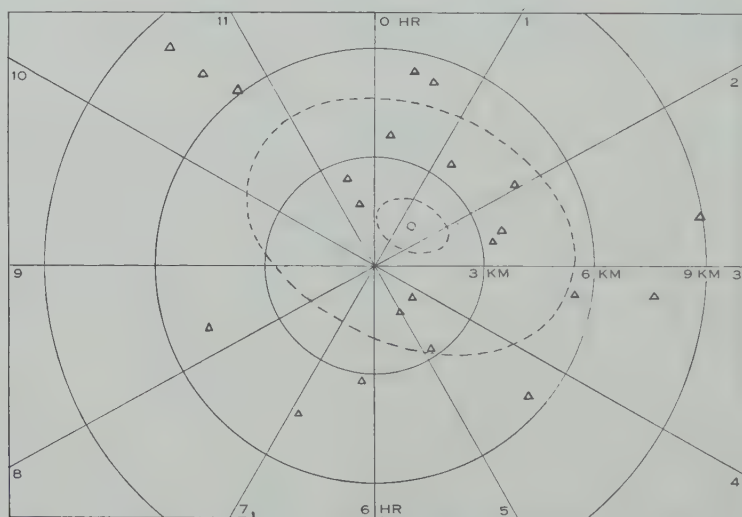


Fig. 1.—Harmonic dial for lunar semi-diurnal variation in $h'E_{2s}$ for Brisbane 1952–1953.

The outer ellipse indicates the probable error of a single month's observation ; the inner one gives the probable error of the mean.

This computed mean has an amplitude of 1.5 km, and epoch of maximum at 1.5 hr. The harmonic dial shows that the mean can be accepted as statistically significant. The marked difference suggested between the tides in E_{2s} and in E_s could be due to the difference in mean equivalent heights, 140 and 110 km respectively in the data analysed.

* Manuscript received December 30, 1957.

† Physics Department, University of Queensland, Brisbane.

‡ THOMAS, J. A., and SVENSON, A. C. (1955).—*Aust. J. Phys.* 8: 554.

PHOTONEUTRONS FROM NATURAL MAGNESIUM*

By B. M. SPICER,† F. R. ALLUM,† J. E. E. BAGLIN,† and H. H. THIES†

Recent experiments on photoneutron production from natural magnesium have produced contradictory results. Katz *et al.* (1954) have used bremsstrahlung from a 22 MeV betatron to irradiate a natural magnesium target. The emitted neutrons were detected with boron trifluoride (BF₃) counters placed in a paraffin block. The counting system was based on a design by Halpern, Mann, and Nathans (1952). The measured yield curve was analysed by the photon difference method (Katz and Cameron 1951*a*) to give a cross section versus energy curve. This indicated two peaks in the magnesium (γ, n) cross section; one of 1.8 mbarn at 13 MeV, and one of 14 mbarn at 20 MeV.

Natural magnesium has three stable isotopes with percentage abundances and (γ, n) thresholds as shown in Table 1. The (γ, n) thresholds were calculated from tables of mass defects given by Wapstra (1955).

TABLE 1
ISOTOPIC ABUNDANCES AND (γ, n) THRESHOLDS OF MAGNESIUM ISOTOPES

Isotope	Abundance (%)	(γ, n) Threshold (MeV)
²⁴ Mg	78.60	16.57
²⁵ Mg	10.11	7.33
²⁶ Mg	11.29	11.12

Katz *et al.* (1954) attributed the 20 MeV peak to the giant resonance of the ²⁴Mg(γ, n) reaction and the 13 MeV peak to the ²⁵Mg(γ, n) giant resonance. The ²⁴Mg(γ, n) ²³Mg reaction had been investigated previously by detecting the 12 sec ²³Mg β^+ -activity (Katz and Cameron 1951*b*). The results obtained in this experiment did not agree with the result obtained from the direct detection of the neutrons (a peak cross section of 9.8 mbarn compared with 14 mbarn). If, however, the residual activity curve is normalized so that the activity measured at 22 MeV is equal to the measured neutron yield from the natural element at 22 MeV, the agreement is much improved.

Nathans and Yergin (1955) have studied the photoneutron yields from enriched isotopes of magnesium (in the chemical form of oxides). Their ²⁴Mg sample was enriched to 99.59 per cent. ²⁴Mg, and their ²⁵Mg sample to 92.33 per cent. ²⁵Mg. Using BF₃ counters contained in a paraffin block they obtained

* Manuscript received January 16, 1958.

† Physics Department, University of Melbourne.

yield curves for the individual magnesium isotopes. The yield curves were corrected for background due to the $^{16}\text{O}(\gamma, n)$ reaction using the results of Ferguson *et al.* (1954), and were analysed for the cross section by the total spectrum method. The $^{24}\text{Mg}(\gamma, n)$ peak cross section (9 mbarn) agreed with the original residual activity result of Katz and Cameron (1951*b*) (9.8 mbarn), but was lower than the renormalized value (14 mbarn). The $^{25}\text{Mg}(\gamma, n)$ cross section had a peak at 20 MeV, which is 7 MeV higher than the result of Katz *et al.* (1954). There was a plateau in the energy range 11–13 MeV.

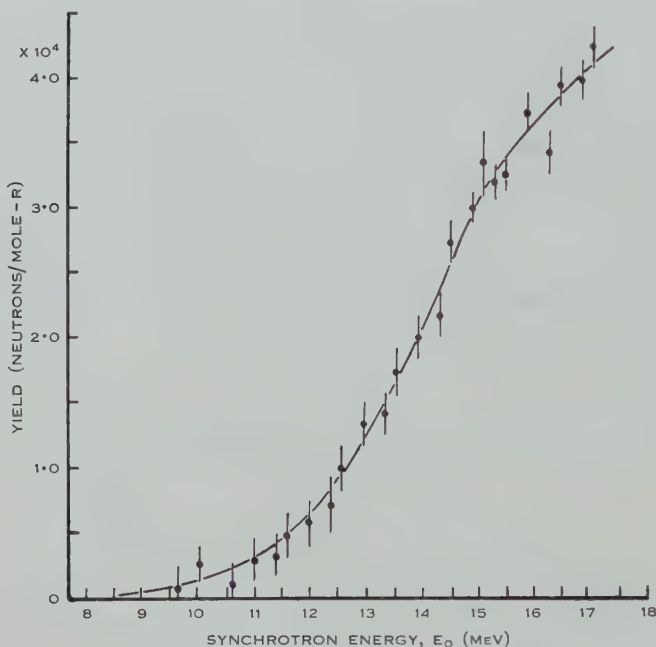


Fig. 1.—Yield of photoneutrons from natural magnesium. Background has been subtracted.

More recently, Yergin (1956) has obtained a neutron yield curve for natural magnesium (using magnesium oxide) which had a different shape from the earlier result of Katz *et al.* (1954). The yield curve obtained here was not analysed to give a cross section. As a result of this measurement, Yergin has said that the results of Nathans and Yergin (1955) cannot be reconciled with the 13 MeV peak obtained by Katz *et al.* He has suggested that Katz's result might be in error if small heavy element impurities were present in the sample.

To provide further experimental evidence, the 18 MeV electron synchrotron at Melbourne has been used to irradiate a natural magnesium target and thus determine the neutron yield as a function of synchrotron energy between 8 and 17 MeV. A sample of the target was analysed spectrographically and found to contain less than 0.1 per cent. aluminium and less than 0.05 per cent. manganese. The photoneutrons were detected by BF_3 counters with essentially the same counting system as in the experiments mentioned above. The one difference is

that no gating circuit was used. This was unnecessary since the γ -ray yield was spread over a 250 μ sec interval, thus removing any danger of a pile-up of electron pulses. The dose was measured by a 0.25 r Victoreen thimble used as a transmission ionization chamber. Its response was compared with that of a 25 r thimble in an 8 cm Lucite cube to find the absolute dose. The neutron yield was corrected for background by subtracting from it the result of a run without the target in position.

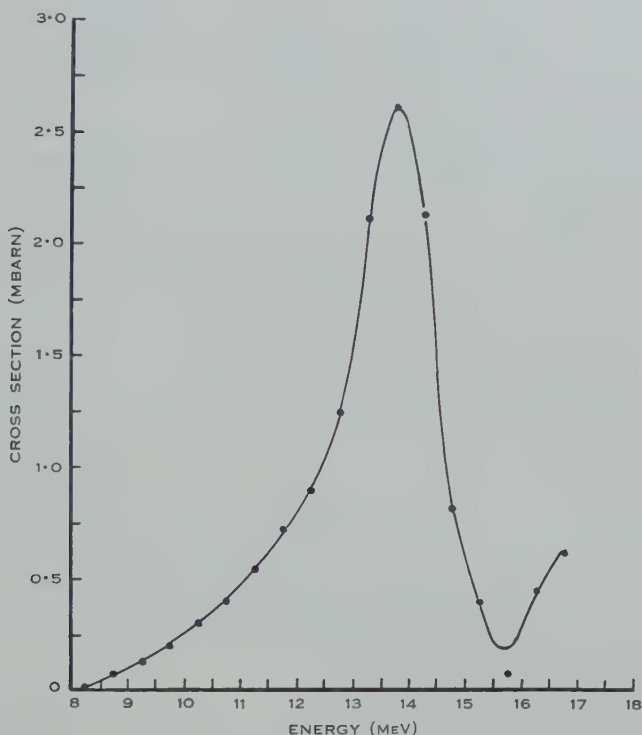


Fig. 2.—Cross section for photoneutron production from natural magnesium.

The resulting yield curve is shown in Figure 1. This curve was analysed for the cross section by the photon difference method after smoothing the first differences of the yield curve as suggested by Katz and Cameron (1951*a*). The cross section curve is shown in Figure 2. It has a peak at 13.5 MeV of height 2.6 mbarn. The absolute yield of neutrons at 17 MeV was calculated to be 4×10^4 neutrons mole⁻¹ r⁻¹, in excellent agreement with the results of the earlier experiments. Katz *et al.* (1954) give 4×10^4 neutrons mole⁻¹ r⁻¹, and Yergin (1956) gives 3×10^4 .

This experiment supports Katz's contention that there is a maximum in the cross section near 13 MeV.

When comparing the yield curve obtained by Katz *et al.* (1954) with his own measurement, Yergin normalized the two curves together at 21 MeV. He then found that they disagreed below 17 MeV. If, however, these yield curves are

normalized together at 15 MeV (below the $^{16}\text{O}(\gamma, n)$ threshold, remembering that Yergin's measurement was done with MgO), the two curves diverge above 15.7 MeV. This is where the neutron yield from oxygen becomes appreciable, suggesting that the oxygen subtraction made by Yergin could be faulty.

In the light of the experiment on the neutron yield from ^{25}Mg by Nathans and Yergin (1955), the original interpretation given to the 13 MeV peak should be revised. This attitude is supported by noting the arbitrary normalization process required to relate the earlier experiments on $^{24}\text{Mg}(\gamma, n)$ using residual activity techniques (Katz and Cameron 1951*b*) with the $^{24}\text{Mg}(\gamma, n)$ cross section deduced from the natural magnesium experiment.

A recent experiment by Cook (1957) has shown the existence of a "pigmy" resonance, that is, a small resonance on the low energy side of the giant resonance, for ^{13}C . With this in mind, we suggest that the peak in the sum of the (γ, n) cross sections for ^{25}Mg and ^{26}Mg may be of this type. The $^{24}\text{Mg}(\gamma, n)$ reaction does not contribute in this energy region because its threshold is at 16.6 MeV. As there has not been an accurate determination of the $^{26}\text{Mg}(\gamma, n)$ cross section, it is not possible to assign the peak to one isotope or the other.

The authors wish to record their appreciation of the interest shown in, and the encouragement given to this work by Professor Sir Leslie Martin.

References

- COOK, B. C. (1957).—*Phys. Rev.* **106**: 300.
FERGUSON, G., HALPERN, J., NATHANS, R., and YERGIN, P. F. (1954).—*Phys. Rev.* **95**: 776.
HALPERN, J., MANN, A. K., and NATHANS, R. (1952).—*Rev. Sci. Instrum.* **23**: 678.
KATZ, L., and CAMERON, A. G. (1951*a*).—*Canad. J. Phys.* **29**: 518.
KATZ, L., and CAMERON, A. G. (1951*b*).—*Phys. Rev.* **84**: 1115.
KATZ, L., HASLAM, R. N., GOLDBERG, J., and TAYLOR, J. G. (1954).—*Canad. J. Phys.* **32**: 580.
NATHANS, R., and YERGIN, P. F. (1955).—*Phys. Rev.* **98**: 1296.
WAPSTRA, A. H. (1955).—*Physica* **21**: 367.
YERGIN, P. F. (1956).—*Phys. Rev.* **104**: 1340.

A SUGGESTED IMPROVEMENT TO THE C.W. TECHNIQUE FOR MEASUREMENT OF METEOR VELOCITIES*

By J. S. MAINSTONE,[†] W. G. ELFORD,[†] and A. A. WEISS[‡]

Radio methods of measuring meteor velocities, whether by pulse (Davies and Ellyett 1949) or C.W. techniques (McKinley 1951), ultimately depend on a knowledge of the separation, along Cornu's spiral, between points representing any two maxima or any two minima of the echo waveform. The relation $s = \frac{1}{2} \nabla \sqrt{(R_0 \lambda)}$, where R_0 is the minimum range of the echo and λ is the wavelength, enables these "zone values" ∇ to be correlated with line segments s of the meteor trail. The time taken for the meteor to traverse the distance s is obtained from the echo, and hence the velocity of the meteor can be found.

Most sources of error in velocity determination can be attributed to distortion of Cornu's spiral in some way or another, and are common to both pulse and C.W. systems of measurement. The effects of wind shear (Kaiser 1955) and non-uniformity of ionization cannot be eliminated, in general, but are not serious. The effect of transverse polarization (Billam and Browne 1956) is less serious in the C.W. technique where the zones normally used are those prior to the t_0 point. Under conditions of severe diffusion (Dr. C. D. Ellyett, personal communication) large errors (up to 20 per cent.) can be introduced into the pulse measurements, but again the error is not serious in the C.W. method if zones of high order prior to the t_0 point are used.

The C.W. method is therefore inherently more accurate but suffers from the grave defect that the phase of the direct wave, to which Cornu's spiral for the reflected wave is referred, must be known. This cannot be inferred from the "whistle" waveform with any certainty. A detailed analysis of the C.W. method suggests that large errors can be introduced by this—up to 10 per cent. in the case of the zone defined by the 4th and 9th maxima used by McKinley—although in general the error is likely to be much less than the maximum possible 10 per cent.

By using the complete C.W. echo waveform, comprising both whistle and body Doppler beat pattern due to the wind drift of the trail, in conjunction with knowledge of the sense of the wind drift, which in our case is conveniently obtained from the phase spikes (Robertson, Liddy, and Elford 1953), it appears possible to determine unambiguously the phase of the direct (reference) wave at the times represented by any maxima of the whistle which are used in the measurement of the meteor velocity. This modification of the C.W. method

* Manuscript received February 6, 1958.

[†] Department of Physics, University of Adelaide.

[‡] Division of Radiophysics, C.S.I.R.O., at Department of Physics, University of Adelaide.

should prove to be of great advantage for the accurate measurement of the velocities of faint meteors where the application of pulse techniques becomes difficult because of the rapid decay of the echo.

Equipment to test the practical application of this refinement of the C.W. technique is at present under construction.

A full account of these investigations will be published elsewhere.

References

- BILLAM, E. R., and BROWNE, I. C. (1956).—*Proc. Phys. Soc. Lond. B* **69**: 98.
DAVIES, J. G., and ELLYETT, C. D. (1949).—*Phil. Mag.* **40**: 614.
KAISER, T. R. (1955).—"Meteors." (Ed. T. R. Kaiser.) p. 55. (Pergamon Press: London.)
McKINLEY, D. W. R. (1951).—*Astrophys. J.* **113**: 225.
ROBERTSON, D. S., LIDDY, D. T., and ELFord, W. G. (1953).—*J. Atmos. Terr. Phys.* **4**: 255.

ETHER AND RELATIVITY

By G. BUILDER*

[Manuscript received November 20, 1957]

Summary

The relative retardation of clocks, predicted by the restricted theory of relativity, demands our recognition of the causal significance of absolute velocities. This demand is also implied by the relativistic equations of electrodynamics and even by the formulation of the restricted theory itself. The observable effects of absolute accelerations and of absolute velocities must be ascribed to interaction of bodies and physical systems with some absolute inertial system. We have no alternative but to identify this absolute system with the universe. Thus, in the context of physics, absolute motion must be understood to mean motion relative to the universe, and any wider or more abstract interpretation of the "absolute" must be denied.

Interaction of bodies and physical systems with the universe cannot be described in terms of Mach's hypothesis, since this is untenable. There is therefore no alternative to the ether hypothesis. It is shown that this is compatible with the restricted theory of relativity and even provides a tenable basis, when taken together with the principle of relativity, for that theory. It is shown that the hypothesis provides a satisfactory and sufficient causal explanation of the predicted relative retardation of clocks, and attention is drawn to its striking pedagogical and heuristic advantages.

I. INTRODUCTION

In a recent letter to *Nature*, Professor Dingle (1957) has pointed out that the relative retardation of clocks, said to be predicted by the restricted theory of relativity (e.g. Einstein 1905; Builder 1957a) is an absolute effect which is a function of the velocities of the clocks. He concludes that: "*It should be obvious that if there is an absolute effect which is a function of velocity then the velocity must be absolute. No manipulation of formulae or devising of ingenious experiments can alter that simple fact.*"

Professor Dingle himself holds that this statement demonstrates that the prediction must be wrong. He claims that the restricted theory of relativity is incompatible with the ascription of causal significance to absolute velocities and, in particular, that it is incompatible with the existence of an ether. It will be shown in the following that this claim cannot be sustained.

The importance of the statement is obvious. It means that, if the prediction is correct, the restricted theory thereby demands our recognition of the causal significance of absolute velocities, even although it requires that we should be unable to detect or measure such velocities by observations of dynamical and electro-dynamical phenomena.

In other words, if two isolated clocks move, in a region of the universe free of gravitational fields, in such a way that they coincide on two or more occasions,

* School of Physics, University of Sydney.

the restricted theory predicts that, in general, one will become retarded relative to the other in the interval between their successive coincidences, i.e. it predicts that, in general, something different will happen to the two clocks as a result of their individual motions. This cannot be related causally to their individual accelerations and must be related causally to their individual velocities.

More precisely, according to the restricted theory, the calculation of the relative retardation of the two clocks requires a knowledge of their individual speeds as measured in some inertial reference system. It is not sufficient to know the velocity of the clocks relative to one another; it would still be necessary to know the velocity of one of the clocks and we would then know the speeds of both. In the much-quoted simple case in which it is postulated that one of the clocks remains at rest in an inertial reference system, so that the velocity of the second clock as measured in this system may be regarded as its velocity relative to the clock at rest, it remains true that the individual speeds of both clocks are thus specified and both are used in the calculation of the relative retardation; this corresponds to putting $u=0$ in the general case given below.

Suppose that, according to the measures of some one inertial reference system S , the speeds of two standard clocks A and B are u and v at any instant t and their coincidences occur at the times t_1 and t_2 . Then, according to measurements made in S , the rates of the clocks at the instant t are, respectively,

$$dt_a/dt = (1 - u^2/c^2)^{\frac{1}{2}}; \quad dt_b/dt = (1 - v^2/c^2)^{\frac{1}{2}}, \quad \dots \dots \dots (1)$$

so that the proper times of the clocks between their coincidences are given respectively by

$$t_a = \int_{t_1}^{t_2} (1 - u^2/c^2)^{\frac{1}{2}} dt; \quad t_b = \int_{t_1}^{t_2} (1 - v^2/c^2)^{\frac{1}{2}} dt. \quad \dots \dots (2)$$

Thus, in the interval between their coincidences, clock B becomes retarded relative to clock A by the amount

$$t_a - t_b = \int_{t_1}^{t_2} (1 - u^2/c^2)^{\frac{1}{2}} dt - \int_{t_1}^{t_2} (1 - v^2/c^2)^{\frac{1}{2}} dt, \quad \dots \dots \dots (3)$$

and it may be shown that the value of this expression is an invariant for *all* systems of reference and for *all* observers.

This expression for the relative retardation is not a function of the accelerations of the clocks; it therefore fails to suggest, and even precludes, the possibility of a causal relation between the relative retardation and these accelerations. This is consistent with the generally accepted view, which is basic in the general theory of relativity, that the rate of a clock is not a function of its acceleration.

Nor can the expression be written as a function of the velocity of one of the clocks relative to the other. Thus it does not suggest, and indeed precludes, any possibility of a causal relation between the relative retardation and this relative velocity. This is consistent with the fact that the context of the problem precludes any physical interaction between the two clocks, and hence precludes

any possibility of ascribing causal significance to the motion of one relative to the other.

The expression of equation (3) is an explicit function of the individual speeds u and v of the clocks. It therefore suggests, and indeed implies, a causal relation between the relative retardation and these individual speeds; this obviously arises out of a causal relation, implied by equations (1), between the rates of the individual clocks and their individual speeds. In other words, equations (1) themselves require that the two clocks behave differently if their speeds measured in S are different, and require also that each clock will change its behaviour when its speed, measured in S , changes.

Any physical explanation of these causal relations must obviously be independent of the inertial reference system chosen for measurement or calculation. This is clearly required by the fact that the equations (1)–(3) take precisely the same form when expressed in the measures of any inertial reference system whatsoever. It is also required by the context, for this precludes any physical interaction of the clocks with any such reference system. This may be illustrated as follows. We could, if we wished, regard u and v in equation (3) as the speeds of the clocks relative to the system S , as measured in S . Yet we could not ascribe direct causal significance to these speeds relative to S , because any corresponding interaction, between the clocks and S , is precluded by the context. Indeed, the equations hold even if the system S is purely hypothetical and if the quantities in the equations are merely postulated, or if they are calculated from measurements made in some other system.

It follows from this that any physical explanation of the phenomena described in equations (1)–(3) must be sought in the form of these equations rather than in their numerical content as determined by the measures of the system S . The fact that the form of the equations is independent of the choice of the inertial reference system implies that the existence in nature of the phenomena described by the equations is independent of the existence of any such inertial reference systems, hypothetical or physical.

Yet the fact that the clocks do behave differently when their speeds are different requires that they interact physically with *something*, in a manner which depends on their speeds. For the context requires that the two clocks be ideal standard clocks which behave identically in all respects when subject to the same conditions. Thus any difference in their behaviour must be ascribed to a difference in their physical interaction with their environment.

Since the context requires that the clocks be isolated from interaction with other actual bodies or physical systems in their vicinity, we are forced to conclude that they must interact with something universal or with the universe as a whole. This conclusion is permissible because the existence of the universe is implied by the context. Indeed, as I have pointed out elsewhere (Builder 1957*b*) the problem being considered would fall outside the domain of physical enquiry if this were not the case; physics can give no guide as to what might be expected to happen to clocks isolated in an abstract empty space not related to this universe.

The only hypothesis that is tenable, and that is compatible with the foregoing considerations, is that there exists a unique absolute inertial system, such as the universe as a whole, which interacts with, and affects the behaviour of, the clocks in a manner dependent on their speeds relative to it, i.e. their absolute speeds.

This hypothesis is clearly sufficient. A reference system S_0 at rest relative to this postulated absolute inertial system would be one of the reference systems to which the restricted theory is applicable. We can therefore write for the rates of the clocks A and B , as measured in S_0 ,

$$dt_a/dt_0 = (1 - u_0^2/c^2)^{\frac{1}{2}}; \quad dt_b/dt_0 = (1 - v_0^2/c^2)^{\frac{1}{2}}, \quad \dots\dots\dots (4)$$

and for the relative retardation,

$$t_a - t_b = \int_{t_{01}}^{t_{02}} (1 - u_0^2/c^2)^{\frac{1}{2}} dt_0 - \int_{t_{01}}^{t_{02}} (1 - v_0^2/c^2)^{\frac{1}{2}} dt_0, \quad \dots\dots\dots (5)$$

where u_0 and v_0 are the absolute speeds of the clocks at each absolute time t_0 .

We thus have in equations (4) and (5) a causal account of the behaviour of the clocks given explicitly in terms of their absolute speeds u_0 and v_0 . It is true that we cannot measure these speeds, because we cannot identify the system S_0 ; but this is not necessary, because all the *observable* consequences of (4) and (5) can be verified by measurements made in any inertial system S and by calculations using equations (1) and (3). In other words, although equations (1) and (3) do not contain u_0 and v_0 explicitly, they do express, in terms of the speeds u and v , all the *observable* consequences of equations (4) and (5).

Thus we conclude that the relative retardation of clocks predicted by the restricted theory does indeed compel us to recognize the causal significance of absolute velocities and that this recognition is compatible with the fact that these absolute velocities do not appear explicitly in the relativistic expression for the relative retardation, except in the unique and unidentifiable case in which the inertial reference system considered is at absolute rest.

The relative retardation of clocks is an effect which seems to be unique in that its measure is an invariant for all observers, whatever their state of motion. However, it is important to realize that this unique character arises solely from the fact that each of the clocks considered in this context incorporates an integrating device which provides an observable record of the accumulated effects of variations in its rate. Were we considering the periodic processes in a single atom, we would be without such a cumulative record; but, as has been indicated above, equations (1) would still require us to postulate some absolute system which would affect the rate of these periodic processes in accordance with the absolute speed of the atom.

The corresponding relativistic variations of the mass and of the dimensions of a body similarly imply the existence of some absolute system which affects the mass and the dimensions of the body in accordance with its absolute speed. There is not, however, available any known mechanism which, like the integrating

mechanism of a clock or the observable senescence of an animal, could provide us with any record of the cumulative effect of such variations. Yet it is clear that, were such mechanisms available, the restricted theory would predict that they would show cumulative effects analogous to the relative clock retardation. *Thus the demand for our recognition of the causal significance of absolute velocities is implied not only in the relativistic variations of the rates of clocks but also in the relativistic variations of the masses and dimensions of bodies.*

It is also implied in the relativistic equations of electrodynamics and even in the context of the restricted theory of relativity itself, as is shown in the next two sections.

II. ELECTRODYNAMICAL PHENOMENA

The relativistic equations of electrodynamics are, in form, identical to the Maxwell-Lorentz equations. They differ in that the velocities occurring in them are defined as the velocities measured in the particular reference system being used, whereas the velocities in the Maxwell-Lorentz equations are defined as being relative to the ether, i.e. as absolute velocities. As is well known, the relativistic equations hold in every inertial reference system as relations between quantities measured in that system.

It follows that electrodynamical phenomena, as observed in any inertial reference system S , will display characteristics which are precisely the same as those that would, according to the Maxwell-Lorentz theory, be displayed in a reference system S_0 at rest in the ether. Thus the asymmetries and the dependence of the phenomena on the individual velocities of particles and bodies, which were such notable features of the Maxwell-Lorentz equations, are retained in the relativistic equations of electrodynamics and must necessarily be displayed in electrodynamical phenomena as observed in any inertial reference system.

It can be demonstrated that the phenomena observed in any inertial reference system S are determined, at least in part, by the individual velocities and accelerations of particles and bodies. The phenomena cannot, in general, be described solely in terms of the velocities and accelerations of the particles and bodies relative to one another, and it will be shown that this is true whether these relative velocities and accelerations are measured in the inertial reference system S (as they should be) or in the rest systems of the particles and bodies considered. It can also be shown that the phenomena observed in S display marked asymmetries which are incompatible with any supposition that the phenomena depend only on the motions of the particles and bodies relative to one another; in particular, Newton's third law does not generally hold.

These points can best be demonstrated by considering the interactions of two point particles. For simplicity, only the interactions which depend on the particle velocities will be considered in detail. This is possible because the main effects of the velocities are independent of the accelerations. In any case, it is already well known and generally appreciated that the phenomena do depend, at least in part, on the individual accelerations of particles as, for example, in the radiation from an isolated point-charge which is subject to acceleration.

Consider first the interaction between a magnetic point-pole m and a point-charge q as observed in any inertial reference system S . It will be sufficient to consider two special cases.

- (a) The pole m at rest at the origin; the charge q moving with uniform velocity \mathbf{v} at the point \mathbf{r} .
- (b) The pole m at the origin moving with the uniform velocity $-\mathbf{v}$; the charge q at rest at the point \mathbf{r} .

The instantaneous positions are the same in the two cases, and the velocity of the charge relative to the pole is, in each case, equal to \mathbf{v} ; this is true whether the relative velocity is measured in S (as it should be) or in the rest system of either of the two particles.

In case (a), with the pole at rest, the force \mathbf{F} experienced by the charge, as measured in S , is given by

$$\mathbf{F} = (1/c)q\mathbf{v} \times \mathbf{B}_0, \dots\dots\dots (6a)$$

where \mathbf{B}_0 is the magnetic flux density, due to the pole, at the position \mathbf{r} of the charge.

In case (b), with the charge at rest, the force \mathbf{F}' experienced by the charge, as measured in S , is given, approximately, by

$$\mathbf{F}' = \frac{1}{c}q\mathbf{v} \times \mathbf{B}_0 \left\{ 1 - \frac{3}{2}(\mathbf{v} \cdot \mathbf{r})^2/r^2c^2 + \frac{1}{2}v^2/c^2 \right\}, \dots\dots\dots (6b)$$

where \mathbf{B}_0 has the same value as above. This equation may be derived by appropriate relativistic transformations, or it may be derived directly from the relativistic equations of electrodynamics, taking into account the retardations of the potentials.

Thus the force experienced by the charge, as measured in S , is different in the two cases, in spite of the fact that the relative velocity is the same. This difference can be ascribed to the finite time required for the transmission of electromagnetic effects, as expressed by the retardation of the potentials. Furthermore, it is easy to show that, in either case, the force experienced by the pole will usually differ from that experienced by the charge, so that Newton's third law will usually not hold.

This particular example is of special interest because of its relation to the phenomenon of electromagnetic induction displayed by a magnet and a conducting circuit when in motion relative to one another. At least in principle, equations (6a) and (6b) would permit the observable induction phenomena to be predicted in any specific case once the magnet and circuit configurations and motions had been prescribed. It follows that the observable phenomena in the case of magnet and conductor must depend in part on the individual velocities of the magnet and conductor. This can readily be verified for the simple case of a magnet moving, in the direction of its length, along the axis of a circular conducting circuit, by taking into account the retardation of the potentials.

It is true that, in the limiting case of velocities very small compared with that of light (so that transmission delays, and the consequent retardation of the

potentials, can be neglected), the effects can be described in terms of the relative velocity alone. In this case equation (6b) is reduced to equation (6a). This may be a perfectly satisfactory approximation, from a practical point of view, in dealing with magnets and conductors under laboratory conditions; but the view expressed by Einstein, in the introductory paragraphs of his 1905 paper, that it is a fundamental characteristic of the phenomena, must be rejected. The fact that this view may have somehow suggested the restricted theory to Einstein is not a tenable argument in its favour, since he did not utilize it as a premise of the theory.

Apart from this special interest, it is better to consider the more general case of the interaction of two point charges q_1 and q_2 , moving with velocities \mathbf{v}_1 and \mathbf{v}_2 at the simultaneous positions \mathbf{r}_1 and \mathbf{r}_2 in S .

It can be shown that the force \mathbf{F}_1 experienced by q_1 , as measured in S , is given, approximately, by

$$\mathbf{F}_1 = \frac{q_1 q_2}{r^2} \left[1 - \frac{3}{2} (\mathbf{v}_2 \cdot \hat{\mathbf{r}})^2 / c^2 + \frac{1}{2} v_2^2 / c^2 \right] \left[\hat{\mathbf{r}} + \frac{1}{c^2} \mathbf{v}_1 \times (\mathbf{v}_2 \times \hat{\mathbf{r}}) \right], \quad \dots (7)$$

where

$$r\mathbf{r} = \mathbf{r}_1 - \mathbf{r}_2.$$

This is an asymmetrical function of the individual velocities \mathbf{v}_1 and \mathbf{v}_2 and cannot be expressed as a function of the relative velocity $\mathbf{v}_1 - \mathbf{v}_2$.

If q_1 is at rest and q_2 is moving with velocity \mathbf{v} , the force experienced by q_1 is given by

$$\mathbf{F}_1 = \frac{q_1 q_2}{r^2} \left[1 - \frac{3}{2} (\mathbf{v} \cdot \hat{\mathbf{r}})^2 / c^2 + \frac{1}{2} v^2 / c^2 \right] \hat{\mathbf{r}}, \quad \dots (7a)$$

whereas, when q_1 is moving with velocity $-\mathbf{v}$ and q_2 is at rest, the force experienced by q_1 is

$$\mathbf{F}_1 = \frac{q_1 q_2}{r^2} \hat{\mathbf{r}}, \quad \dots (7b)$$

Thus the force experienced by q_1 is different in the two cases. Yet the velocity \mathbf{v} of q_2 relative to q_1 remains the same, and this is so whether it is measured in S (as it should be) or in the rest system of either point charge.

Furthermore the force \mathbf{F}_1 experienced by the charge q_1 is not equal and opposite to the force \mathbf{F}_2 experienced by q_2 which, in the general case, is given by

$$\mathbf{F}_2 = - \frac{q_1 q_2}{r^2} \left[1 - \frac{3}{2} (\mathbf{v}_1 \cdot \hat{\mathbf{r}})^2 / c^2 + \frac{1}{2} v_1^2 / c^2 \right] \left[\hat{\mathbf{r}} + \frac{1}{c^2} \mathbf{v}_2 \times (\mathbf{v}_1 \times \hat{\mathbf{r}}) \right]. \quad \dots (8)$$

This asymmetry, between the forces experienced simultaneously by the two charges, obviously precludes any inference that the forces are determined solely by the relative motions of the charges.

At least in principle, all electromagnetic phenomena can be described in terms of the force formulae of equations (7) and (8). It is by no means difficult to derive in this way an expression for the force between two elements of current-

carrying conductors, or the force exerted by a current element on a moving charge. By such calculations, it may be shown that the force exerted by one current element on another is a function of the velocities of the charges within each current element and is not, even in part, determined by the velocities of the charges in either element relative to the charges in the other; and, as is well known, the forces between the elements are asymmetrical and do not satisfy Newton's third law. Similarly it may be shown that, as is well known, the force exerted by a current element on a charge q is a function of the measured velocity \mathbf{v} of the charge and cannot be expressed as a function of its velocity relative to the current element as a whole, or relative to the charges in the current element.

Thus observers in any initial reference system S , and therefore also observers in any inertial reference system whatsoever, must agree

- (i) that the electrodynamical phenomena are determined, at least in part, by the individual velocities of particles and bodies, and
- (ii) that the observed phenomena are characterized by asymmetries which are incompatible with any supposition that they are determined solely by the motions of the particles relative to one another.

Thus the relativistic equations of electrodynamics imply causal relations between the phenomena observable in any inertial reference system, whether actual or hypothetical, and the individual velocities measured in that same system.

As in Section I, the context precludes any possibility of a physical explanation of these causal relations in terms of interaction of the particles and bodies with any one arbitrarily selected inertial reference system. Moreover, since the equations have the same form for every such inertial reference system, actual or hypothetical, the physical explanation of the implied causal relations must be sought in the form of the equations rather than in the numerical content of the equations in any arbitrarily selected inertial reference system. The required physical explanation is immediately available. *It is only necessary to postulate that the phenomena are caused by motions of particles and bodies relative to an absolute inertial system in accordance with the Maxwell-Lorentz equations.*

That this postulate is permissible in the context of the restricted theory is obvious. For an inertial reference system S_0 at rest relative to such an absolute inertial system is one of the permissible reference systems of the restricted theory. Thus the relativistic equations of electrodynamics must hold in the system S_0 ; these are identical in form with the Maxwell-Lorentz equations and become synonymous with them for the system S_0 in which the measured velocities are absolute velocities.

That this postulate is sufficient is also obvious. In the Maxwell-Lorentz equations the electrodynamical effects are related explicitly to the absolute velocities which cause them. It is true that we cannot measure these absolute velocities; but, subject to this limitation, all observable predictions of the Maxwell-Lorentz equations can be verified by observations made in any arbitrarily selected inertial reference system. This may be expressed more specifically by saying that the postulate is sufficient because, if it is correct, it follows that the

relativistic equations of electrodynamics must hold in any arbitrarily selected inertial reference system.

Moreover, the postulate is necessary, because there is not available any other tenable physical explanation of the causal relations, between electrodynamical phenomena and the individual velocities of particles and bodies, implied by the relativistic equations of electrodynamics.

Thus the observable characteristics of electrodynamical phenomena, like the observable behaviour of clocks, demands our recognition of the causal significance of absolute velocities.

III. THE RESTRICTED THEORY OF RELATIVITY

The restricted theory of relativity states that the spatial and temporal coordinates of an event, measured in any one inertial reference system, are related by the Lorentz transformations to the spatial and temporal coordinates of the same event, measured in any other inertial reference system.

The theory is thus restricted to measurements made in inertial reference systems. It is, of course, quite inapplicable to measurements made in systems in uniform motion relative to one another unless each of these systems is itself inertial. An inertial system is one in which the Newtonian laws of mechanics hold to a first approximation (Einstein 1905). This means that, if it is in a region of the universe free of gravitational fields, the system must be unaccelerated.

It is generally recognized and accepted that the acceleration of a physical system which is in a region of the universe free of gravitational fields can be detected by observations, within the system, of dynamical phenomena. If this were not so, it would be impossible to identify the class of reference systems, i.e. the inertial ones, to which the restricted theory is applicable.

The absolute character of acceleration forces us, like Newton, to postulate some absolute universal system relative to which bodies and systems are accelerated. Moreover, the possibility of detecting absolute acceleration by its dynamical effects forces us to ascribe these effects to interaction between the bodies affected and this absolute system. In particular, it is necessary to ascribe the inertia of bodies to such interaction. The necessity for this has been argued fully by Mach, Einstein, and others.

This absolute system, implied by the absolute character of accelerations, is of necessity itself an inertial system. It is therefore one of the infinity of inertial systems to which the restricted theory is, in principle, applicable. However, it is distinguished from all the inertial reference systems of the restricted theory by the fact that it alone interacts physically with bodies and systems. *Thus the possibility of defining and identifying the class of inertial reference systems to which the restricted theory is applicable demands the recognition that there must exist a unique absolute system which is itself inertial and which interacts physically with bodies and systems to cause the dynamical effects which enable us to detect accelerations relative to it.*

Acceleration of a body relative to this absolute system must, by the definition of acceleration, result in a change in its absolute velocity. Thus the existence in

nature of something corresponding to the concept of absolute acceleration *entails* the existence in nature of something corresponding to the concept of absolute velocity.

It follows that any body or system which is accelerated must thereby suffer a change in its absolute velocity. Similarly, any two bodies which are in motion relative to one another must be moving with different absolute velocities. These statements are necessarily true even though we cannot measure the individual absolute velocities.* We can therefore determine whether the absolute velocity of a body remains constant or changes, and we can determine whether two bodies have the same or different absolute velocities. Moreover, these determinations can be made by purely kinematical observations.

It is true that, in accordance with the *principle of relativity*, we cannot measure or detect the absolute velocities of bodies by observations of dynamical or electrodynamical phenomena. Yet it cannot be validly inferred from this that absolute velocities of bodies could not be measured by other means; this point is discussed further in Section IV below. Nor can the principle of relativity preclude the possibility that there might be observable differences between the behaviour or properties of two bodies whose absolute motions are different and it has been shown, in Sections I and II above, that such differences are in fact observable. Since we have already been forced to postulate an absolute system which interacts with bodies to cause the observable effects of their absolute accelerations, it is reasonable to suppose that the same absolute system may interact with the bodies to cause observable differences, in their behaviour and properties, corresponding to the fact that their absolute velocities are different.

This thesis can be well illustrated by the restricted theory itself. Let there be supposed to exist two inertial reference systems S and S' , similarly equipped for the making of physical measurements, which would in all respects be identical if compared together at rest. In other words, the systems are supposed to differ only in that their uniform motions are different. The restricted theory requires that measurements made in S and S' of the behaviour and properties of some other physical system F should give different results. For example, the mass of a particular body in F , the length of a particular rod in F , or the rate of a particular clock in F , will be different according to the measures of S and S' .

It is true that we can predict the relation between these measures in S and S' if we know the velocity of S' relative to S . Yet we cannot ascribe any causal significance to this relative velocity because, in the context of the restricted theory, interaction between the systems S and S' is precluded by hypothesis. The context of the restricted theory similarly precludes the possibility of physical interaction between the systems S and F , or between S' and F .

* This argument has been set out in much greater detail by Wiechert (1911), who coined the word "Schreitung" to denote the condition of a body corresponding to what we have here referred to as its absolute velocity. A special term has the advantage that its use precludes any misunderstandings which might arise through attaching to the much-used word "absolute" implications lying outside the scope of the above discussion.

Yet the measures of S and S' are in fact different. There remains no alternative to assuming that this difference is caused by the difference in the absolute velocities of S and S' .

The correctness of this assumption is strongly supported by our knowledge that measurements made in any inertial reference system show that the behaviour of clocks and of measuring rods differ if their measured speeds differ, i.e. if their absolute speeds differ. We are thereby compelled to infer that the behaviour of the clocks and measuring rods used in S and S' must be different. Such a difference in behaviour of the devices used for measurement must necessarily lead to, and can be shown to provide a detailed explanation of, the difference in the measures of S and S' .

Thus the restricted theory itself requires our recognition of the existence of an absolute inertial system which interacts with bodies and physical systems in a manner depending on their absolute velocities. In other words, it requires our recognition of the causal significance of absolute velocities.

IV. ABSOLUTE SPACE AND MOTION

It has been shown in the foregoing that the characteristics of physical phenomena, and the accepted formulation of the restricted theory of relativity, are such that we are compelled to admit the existence of an absolute inertial system which interacts physically with bodies and physical systems in a manner which depends on their accelerations and their velocities relative to it.

There is no feasible alternative to supposing that this absolute system is the universe as a whole, or else something universal which is an integral and essential part of the universe. *The term "absolute" must therefore be understood here to characterize anything which is defined, or measured, relative to the universe.* To accept any more abstract interpretation of the absolute would be meaningless in the context of physics, since all our physical measurements and all our physical theories relate to this universe and to it alone. Any question as to whether the universe itself is at rest, or in motion, in some broader or more abstract "absolute" sense lies outside the domain of physical enquiry.*

Thus, in this present context, *absolute motion* is to be understood as motion relative to the universe as a whole and *absolute space* as space defined by the

* This is the basis of the dictum of Poincaré (1908): "Whoever speaks of absolute space uses a word devoid of meaning." He was, of course, then using the term in an abstract sense not uncommon in philosophy and he was, in fact, distinguishing this abstract absolute space from a universal ether at rest in the universe. This distinction has unfortunately sometimes been overlooked; for example, Jammer (1954, p. 142) wrongly treats Poincaré's concepts of absolute space and of the ether as being identical and, incorrectly, quotes Poincaré's dictum as applying to both.

It is to be noted that our definition of the "absolute" cuts across the distinction, much discussed by philosophers, between Absolute and Relational Theories of Motion (e.g. see Broad 1923). Yet if we are to continue to use the word "absolute" at all in physics, we cannot ascribe any useful meaning to it other than that adopted here. As a philosopher, Professor Alan Stout's reference to this as the "relatively absolute", though perhaps not very seriously intended, is not without point.

universe as a whole. Instead of the "universe as a whole" we can simply speak of the "universe" or, since the universe is primarily characterized by the distribution of matter, we could instead speak of the "distributed masses of the universe" as representing the universe approximately. As a still cruder approximation, we may usefully and picturesquely speak of the "fixed" stars as representing the universe, even though these stars are known to be in motion relative to it.

It must be recognized that these definitions of absolute space and of absolute motion are, respectively, purely geometrical and purely kinematical. They imply nothing whatsoever about the dynamical aspects of motion or about the physical characteristics of space. They do, however, make definite the concepts of absolute acceleration and of absolute velocity and enable us to prescribe how these are to be measured, i.e. relative to the universe or, approximately, relative to the fixed stars.

In practice, we can in fact make such purely kinematical measurements of absolute rotation; we need only observe the rotation, of axes fixed in a body, relative to the fixed stars. On the other hand, purely kinematical measurements of small absolute linear accelerations and of small absolute velocities seem to be quite impracticable. The limitations in the accuracy of our measurements, the lack of any rigid reference framework of sufficient extent, and the time required by light to traverse the vast spaces of the universe, all preclude our mapping the absolute space of the universe with sufficient accuracy to measure small linear accelerations, or small velocities, relative to it. Even if these limitations were reduced, it would still seem to be impossible to deal with the very complex distribution and motion of matter in the universe.

In spite of the nature and extent of these limitations, it seems proper to regard them as being essentially practical rather than fundamental. They are, in fact, strictly applicable only to small accelerations and velocities. For example, it is generally accepted that the absolute velocity of the solar system cannot be very great; if it were an appreciable fraction of the velocity of light there seems little doubt that this would be revealed by the asymmetry of the Doppler shifts in star light received from different directions.

Thus we are bound to recognize the admissibility of the concept of absolute velocity, and it may even be maintained that such velocities are, in principle, measurable by kinematical methods.

This is not incompatible with the principle of relativity of the restricted theory. The principle requires only that the laws which describe physical phenomena should take the same form for every inertial reference system (Poincaré 1904; Einstein 1905). Thus it requires only that we should be unable to detect absolute uniform motion by observations of dynamical and electrodynamical phenomena. The principle has no relevance whatsoever to the question of whether or not it is possible, in principle, to detect or measure absolute velocities by the purely kinematical methods referred to above.

Nor does the principle of relativity require that there should be no observable effects of the absolute velocities of bodies and physical systems. It requires only that any observable effects must be such that they fail to provide any

measure of absolute velocity. Indeed, it has been shown above that the predictions of the restricted theory compel us to recognize that there are such observable effects.

V. THE ETHER HYPOTHESIS

Thus it has been shown that the dynamical and electrodynamical effects, both of absolute accelerations and of absolute velocities, are observable and must be ascribed to interaction of the affected bodies and physical systems with the universe.

To account for the absolute character of acceleration, Newton had postulated an "absolute space". Mach, like Poincaré, interpreted this concept of Newton in an abstract sense and rejected it as inadmissible in the context of physics.* He postulated (1883) instead that the observed behaviour of bodies is determined by the distributed matter of the universe. This is well known as *Mach's hypothesis*† and it implies that the dynamical behaviour of bodies is determined by their direct and instantaneous interaction with the distributed matter of the universe. Although Mach was dealing only with the observable effects of absolute accelerations, it is obvious that his hypothesis would serve equally to account for any observable effects of absolute velocities.

Mach's *hypothesis* presupposes instantaneous interaction at a distance and cannot now be seriously considered (cf. Einstein 1920, 1924): it cannot be reconciled with the restricted theory of relativity and could in any case be reconciled with the continuous-action field theories of modern physics only by means of a theory including advanced potentials, and no satisfactory theory of this type has yet been, or seems likely to be, worked out.

Thus *Mach's hypothesis must be rejected and there is then no tenable alternative to the ether hypothesis*, i.e. that the space of the universe is endowed with important physical properties and plays a causal role, equal to that played by matter, in physical phenomena. Whether these properties are determined by the distributed matter of the universe, as in Einstein's cosmology, or whether they are *sui generis*, need not concern us here. It does not seem to be generally appreciated that Mach himself (1883, p. 283 of final 1933 edition) discussed the hypothesis of a medium, i.e. what we have called the ether, filling all space and interacting contiguously with bodies to cause the observable effects of acceleration. He held that such an hypothesis is admissible and recognized that it would be sufficient. He deemed his own hypothesis to be more "expedient provisionally"

* Einstein seems never, at least after about 1915, to have seriously considered any such abstract interpretation of Newton's concept, or to have supposed that it could mean anything but the space of the universe endowed with physical properties of causal significance (e.g. Einstein 1920, 1924). Instead, Einstein uses the term absolute, in connexion with space, only to mean that its physical properties are not affected by anything whatsoever, e.g. not even by the presence of matter.

† This is to be distinguished from what Einstein has called *Mach's principle* according to which he utilizes Mach's general idea but supposes that the distributed masses of the universe determine the dynamical behaviour of bodies by determining the properties of space (i.e. of the ether) in the vicinity of the bodies.

only because he claimed that there was then no experimental evidence to determine whether the part played by the distributed masses of the universe, i.e. as an absolute inertial system, is "fundamental or collateral".

The tenability and sufficiency of the ether hypothesis are beyond question. If one is familiar with the historical background of the restricted theory (e.g. Whittaker 1953), with the expositions of the theory presented in the first quarter of the century, and with the whole of Einstein's writings, it is difficult to understand the present widespread view that the ether hypothesis is incompatible with the theory.

There is little doubt that this view originated in Einstein's statement (1905) that: "*The introduction of the 'luminiferous ether' will prove to be superfluous inasmuch as the view to be developed will not require an 'absolute stationary space' provided with special properties, or to assign a velocity vector to a point of the empty space in which electromagnetic processes take place.*" Any doubt that Einstein did, at that time, believe that the ether hypothesis should be discarded is removed by his more definite statement (1907) that: "*. . . the idea of a light ether as a bearer of electric and magnetic fields would not fit into this picture; electromagnetic fields appear here, not as a condition of any medium, but as self-existing things of the same sort as ponderable material and, together with it, have the same characteristic of inertia.*"

Einstein did not present any logical or evidential basis for these opinions. He was dealing with the relations that must subsist between measurements made in different inertial reference systems if the two postulates (that the principle of relativity of uniform motions is valid, and that the velocity of light is independent of the motion of its source) are to be reconciled. There was no *a priori* basis for these postulates; they were generalizations based on experience. Moreover, he adopted a very definite operational approach to the problem of measurement. Thus he could properly claim that he had demonstrated the necessity for adopting the Lorentz transformations without needing to seek any causal explanation of the state of affairs described by the postulates, and so without needing to discuss the ether hypothesis. He could not properly claim any more than this; nor could he, in his chosen context, properly infer anything about the tenability of the ether hypothesis.

It is in any case quite clear that Einstein did not long adhere to the opinions he had expressed in 1905 and 1907. This is clearly demonstrated in his later writings. There is little doubt that this change was forced on him by his formulation of the general theory; he seems first to have stated it in terms of the "ether" when he wrote (1920): "*The next position which it was possible to take up . . . appeared to be the following. The ether does not exist at all. The electromagnetic fields are not states of a medium . . . but are independent realities . . . More careful reflection teaches us, however, that the special theory of relativity does not compel us to deny the existence of an ether . . . On the other hand, there is weighty evidence in favour of the ether hypothesis. To deny the ether is ultimately to assume that empty space has no physical qualities whatever. The fundamental facts of mechanics do not harmonise with this view. According to the general theory of relativity space without ether is unthinkable; for in such space there not only would be no propagation*

of light, but also no possibility of existence for standards of space and time . . . nor therefore any space-time intervals in the physical sense."

Fortunately, there is no ambiguity or uncertainty as to the sense in which Einstein used the term "ether" in these passages. In the same context (1920) he relates the ether concept of which he is speaking to the ether of Lorentz, thus: "*What is fundamentally new in the ether of the general theory of relativity as opposed to the ether of Lorentz consists in this, that the state of the former is at every place determined by connections with the matter and the state of the ether in neighbouring places . . . whereas the state of the Lorentzian ether in the absence of electromagnetic fields is conditioned by nothing outside itself, and is everywhere the same. The ether of the general theory of relativity is transmuted conceptually into the ether of Lorentz if we substitute constants for the functions of space which describe the former, disregarding the causes which condition its state.*" His concept of the ether, at this time, is even more specifically set out, in more technical language and in more detail, in his paper "Über den Äther" (1924); unfortunately there does not seem to be any published English translation of this paper. The subsequent developments of his views are not directly relevant to the present discussion: they are beautifully summarized in a document (1952) which is published as Appendix V in the 1955 edition of his "Relativity".

Quite apart from Einstein's views, we cannot ignore the fact that the restricted theory of relativity had been developed independently by Poincaré and Lorentz before it was presented by Einstein, from a more deductive point of view, in 1905. The historical record, which has been set out in some detail by Whittaker (1953), is quite clear and is readily verified by direct reference to the original literature.

In effect, Poincaré and Lorentz had, in their formulation of the restricted theory, succeeded in reconciling the principle of relativity of uniform motions with the causal significance of velocities relative to the ether, inherent in the Maxwell-Lorentz theory of electrodynamics. In achieving this they did not in any way modify the concept of the ether specified by Lorentz in his extension of the Maxwell theory; but, like Einstein, they were forced to modify the Newtonian mechanics. *Although their derivation does not prove that the ether hypothesis is necessarily correct, it does prove that it is compatible with the restricted theory.*

The compatibility of the ether hypothesis with the restricted theory is also demonstrated in some of the earlier presentations and discussions of the theory. For example, Eddington (1920) presented the theory specifically in terms of the ether hypothesis. His derivation of the Lorentz transformation differs from that of Einstein only in that he starts with a hypothetical system at rest in the ether, whereas Einstein (1905) starts with a system he describes as "stationary". Formally the derivations are identical and that of Einstein is in no way affected if we suppose his "stationary" frame to be also at rest in the ether. Moreover, Einstein's derivation would be in no way affected if his second postulate were reworded in the form: "That light behaves as if it were propagated in a medium, i.e. its velocity does not depend on that of its source."

It is therefore to be concluded that we are without any tenable alternative to the ether hypothesis, and that this hypothesis is not only compatible with the restricted theory but is also a sufficient basis for the theory, i.e. when taken together with the principle of relativity.

It remains only to demonstrate explicitly that the ether hypothesis provides a satisfactory causal explanation of the relative retardation of clocks, and to point out some of its pedagogical and heuristic advantages.

VI. THE RELATIVE RETARDATION OF CLOCKS

The ether hypothesis provides a satisfactory and sufficient causal explanation of the relative retardation of clocks. This claim has already been justified in the discussion of Section I, since the foregoing shows that we must identify the speeds u_0 and v_0 of equations (4) and (5) with the speeds of the clocks relative to the ether. Nevertheless, it seems desirable to set out this causal explanation more briefly and directly as in the following.

If the ether does in fact exist, any reference system S_0 at rest in it is one of the infinity of possible inertial reference systems of the restricted theory. We can therefore use the restricted theory to predict with certainty that a clock moving with speed v_0 relative to S_0 , and hence relative to the ether, will suffer a reduction of its rate by the factor $(1 - v_0^2/c^2)^{\frac{1}{2}}$ compared with the rate of a clock at rest in the ether.

Thus the assumption that motion of a clock through the ether, with speed v_0 , *causes* a change in its rate by the factor $(1 - v_0^2/c^2)^{\frac{1}{2}}$ is compatible with the restricted theory.

On the other hand, this assumption is also consistent with what the Maxwell-Lorentz equations for the ether would lead us to expect. Lorentz (1895) showed that, if the bonds holding the particles of a body together are electromagnetic, or have similar characteristics, then all bodies would be expected to suffer a contraction by a factor $(1 - v_0^2/c^2)^{\frac{1}{2}}$, in the direction of their motion, when moving with speed v_0 through the ether (Fitzgerald-Lorentz contraction). This in turn requires that a clock in motion through the ether must be slowed down* by the factor $(1 - v_0^2/c^2)^{\frac{1}{2}}$, as was first shown by Larmor (1900). If, on the basis of the Maxwell-Lorentz equations, we suppose that the motions of bodies and clocks through the ether do *cause* such reductions in length and in rate, respectively, and if we recognize that true one-way measurements of light velocity are impossible (so that we can measure only the average value of the light velocity over go-and-return paths), it *follows* that measurements made in different inertial

* This immediately follows if one considers a clock consisting of a rigid rod having mirrors at each end to reflect a beam of light backwards and forwards along the length of the rod. If this clock is set in motion with velocity v_0 , it is easy to show that the frequency with which the light traverses the length of the rod is reduced by the factor $(1 - v_0^2/c^2)^{\frac{1}{2}}$. The length contraction thus entails the clock-rate reduction, and the latter is not an independent hypothesis. This point has been overlooked by Broad (1923) and others when they have claimed that the ether theory requires these "two independent hypotheses" to account for the negative result of the Michelson-Morley experiment and for the measured constancy of the velocity of light.

reference systems must be related by the Lorentz transformations. In other words, it follows that the restricted theory of relativity must be valid. This was the course which led to the development of the theory by Poincaré and Lorentz (e.g. Whittaker 1953).

The extensive experimental evidence for the validity of the restricted theory is thus direct evidence for the *tenability* of the assumption that the reductions in length and in clock rates are caused by motion through the ether. This does not *prove* that the assumption is correct or that it is the only possible causal explanation of the fact that measurements made in different inertial reference systems are related by the Lorentz transformations. It does, however, show that the tenability of the assumption can be refuted only by showing its incompatibility with some other part of our physical knowledge of the universe. No grounds for such refutation have yet been demonstrated, and no satisfactory alternative causal explanation has been offered.

We are therefore justified in assuming that a clock in motion through the ether with velocity v_0 experiences a reduction in its rate by the factor $(1 - v_0^2/c^2)^{\frac{1}{2}}$. Then, if two clocks move differently relative to the ether but coincide on successive occasions, it is apparent that, in general, one will become retarded relative to the other during the interval between such coincidences. If, for example, the first of the clocks remains at rest in the ether, and the second makes a journey away from and back to the first, it is obvious that the second must become retarded relative to the first. It is also easily shown that, if the first clock is in uniform motion relative to the ether and the second moves so that it travels away from the first and back to it, then the second must become retarded relative to the first. In both these simple cases, the second clock is observationally distinguishable from the first by the fact that its journey, away from and back to the first, requires that it should be subject to accelerations, whereas the first clock remains free from acceleration.

Now it is true that we cannot ascertain whether a clock is at rest relative to the ether; but we can ascertain that it is at rest or in uniform motion relative to the ether; this requires only that it be at rest or in uniform motion relative to any one inertial reference system. Similarly, we can always ascertain whether a clock is accelerated relative to the ether; this requires only that its motion be accelerated relative to any one inertial reference system (e.g. Builder 1957b).

Thus, if two clocks are observed in an inertial reference system, the first of which is in uniform motion, or at rest, and the second of which is subjected to such accelerations that it moves away from the first and later returns to it, it follows that we must expect the second clock to have become retarded relative to the first in the interval between their coincidences.

Thus the assumption that clocks are slowed down by motion through the ether is a *sufficient* basis for predicting the relative retardation of two clocks which move, in the manner specified, relative to any inertial reference system. This is, of course, identical with the prediction of the restricted theory. We thus provide a tenable causal explanation of the relative retardation of clocks predicted by the restricted theory. In doing so we have ascribed causal significance to the absolute velocities of the clocks, i.e. to their velocities relative to the ether.

VII. CONCLUSION

The permissibility of retaining the concepts of absolute space, of absolute motion, and of the ether, and the fact that we can assign to these concepts definite and clear meanings compatible with the restricted theory of relativity, has striking pedagogical and heuristic advantages.

The conceptual difficulties associated with the restricted theory all arise out of the denial that these absolute concepts are permissible, and out of the consequent attempts to avoid them in the presentation of the theory. It is frequently maintained that the theory has forced us to discard entirely the old-fashioned commonsense notions of time and space; but nothing comprehensible or definable has been offered in their place. Moreover, any questions as to what *causes* the relativity of simultaneity, the measured constancy of the velocity of light in all inertial reference systems, or the reciprocity of relativistic variations of length, of mass, and of clock rates, are evaded by vague references to the principle of relativity, to the four-dimensional character of space-time, and so on.

On the other hand, the presentation of the restricted theory in terms of the absolute concepts (following generally the lines of its development by Poincaré and Lorentz) involves no conceptual difficulties. The relativity of simultaneity, the reciprocity of relativistic variations, and the constancy of the measured velocity of light, then all appear simply as comprehensible effects of the motions, relative to the ether, of the bodies observed and of the measuring instruments used. The only other important factor that has to be taken into account is the impossibility of making true measurements of light velocity over a one-way path, so that we are restricted to measurements which give only an average value of the velocity of light transmitted over a go-and-return path. On this basis the presentation involves no more than a few simple calculations, and can be made easily intelligible and comprehensible to junior university students.

The heuristic value of this approach is also noteworthy. It reduces many questions, which would otherwise lead to discursive and inconclusive arguments, to a form in which a simple and conclusive answer can be given. For example, the relative retardation of clocks predicted by the restricted theory becomes a simple and intelligible consequence of the motion of the clocks relative to the ether.

Similarly, we are enabled to answer intelligibly what are otherwise very difficult and contentious questions, namely: Are the observed relativistic variations of length, of mass, and of clock rates, *real*? How can "reality" be reconciled with their reciprocal character? The answer is, of course, simply that these effects are *really-observable*, but that what corresponds to them in nature are real effects caused by the motions, both of the observed bodies and of the observing instruments, relative to the ether.

It is worth remarking that the pedagogical and heuristic advantages of this approach depend only on the *tenability* of the ether hypothesis and on the admissibility of the absolute concepts, i.e. on their compatibility with the restricted theory and with the general body of physical knowledge. These advantages

would remain even if there were also available an alternative and equally tenable set of hypotheses and concepts.

It has not been practicable to give detailed references to the work of the many previous authors like Wiechert, Lodge, Ives, who have put forward important arguments in favour of the absolute concepts. Still less would it have been practicable to refer to, or to answer, the innumerable ebullient attacks made on the ether hypothesis in the last 50 years.

I have, in fact, tried to confine myself to material relevant to the discussion of what seems to me to be a new, and perhaps decisive, consideration in favour of the absolute concepts, i.e. the recognition (Dingle 1957) that the relative retardation of clocks, predicted by the restricted theory, is an absolute effect which demands our recognition of the causal significance of absolute velocities.

VIII. REFERENCES

- BROAD, C. D. (1923).—"Scientific Thought." (Kegan Paul: London.)
- BUILDER, G. (1957*a*).—The resolution of the clock paradox. *Aust. J. Phys.* **10**: 246.
- BUILDER, G. (1957*b*).—The clock retardation problem. *Aust. J. Phys.* **10**: 424.
- DINGLE, H. (1957).—The "Clock Paradox" of relativity. *Nature* **179**: 866.
- EDDINGTON, A. S. (1920).—"Report on the Relativity Theory of Gravitation." The Physical Society of London. (Fleetway Press: London.)
- EINSTEIN, A. (1905).—Elektrodynamik bewegter Körper. *Ann. Phys., Lpz.* (4) **17**: 891.
- EINSTEIN, A. (1907).—Über das Relativitätsprinzip. . . . *Jb. Radioakt.* **4**: 411.
- EINSTEIN, A. (1920).—"Sidelights on Relativity." (Translation: 1922.) (Methuen: London.)
- EINSTEIN, A. (1924).—Über den Äther. *Verh. Schweiz. Naturf. Ges.* **105**: 85.
- EINSTEIN, A. (1952).—"Relativity." Appendix V (1955). (Methuen: London.)
- JAMMER, M. (1954).—"Concepts of Space." (Harvard Univ. Press.)
- MACH, E. (1883).—"Science of Mechanics." (Ninth (final) Ed. 1933.) (Open Court Publishing Co.: London.)
- LARMOR, J. (1900).—"Aether and Matter." (Cambridge Univ. Press.)
- LORENTZ, H. A. (1895).—"Versuch einer Theorie der elektrischen und optischen Erscheinungen in bewegten Körpern." (E. J. Brill: Leiden.)
- POINCARÉ, H. (1904).—L'état actuelle et l'avenir de la physique mathématique. *Bull. Sci. Math.* **28**: 302. (English translation. *Monist* **15**: 1 (1905).)
- POINCARÉ, H. (1908).—"Science et Méthode." (Trans. Dover Publications 1952.) (Flammarion: Paris.)
- WHITTAKER, E. (1953).—"History of the Theories of Aether and Electricity: 1900–26." (Thomas Nelson and Sons: London.)
- WIECHERT, E. (1911).—Relativitätsprinzip und Äther. *Phys. Z.* **12**: 689.

THE GIANT RESONANCE OF PHOTODISINTEGRATION OF TANTALUM

By B. M. SPICER,* H. H. THIES,* J. E. BAGLIN,* and F. R. ALLUM*

[*Manuscript received April 16, 1958*]

Summary

The neutron yield from tantalum has been measured as a function of bremsstrahlung energy, and from this the shape of the cross section for the nuclear absorption of photons by tantalum has been deduced. This cross section shows definite evidence that it is made up of two component parts. The result is in agreement with the theoretical predictions of Danos (1958) for this nucleus, both regarding the existence of the two components and the energy separation of the two peaks. Using the Danos theory, the intrinsic electric quadrupole moment for tantalum is calculated to be $6.1 \text{ barns} \pm 10 \text{ per cent}$.

I. INTRODUCTION

The theories put forward to account for the nuclear giant dipole resonance fall into two general groups, characterized by the particular type of nuclear model considered.

The independent particle model has been the subject of investigations by Levinger and Bethe (1950), Burkhardt (1953), and Levinger and Kent (1954). Also, Wilkinson (1956) has considered the absorption of radiation by transitions between single-particle states in the shell model. His results are, in general, roughly compatible with observations of the main features of the giant resonance (maximum cross section σ_m , resonance energy E_m , integrated cross section $\int \sigma(E) dE$).

On the other hand, the model assuming long-range correlations between nucleons in the excited state, or hydrodynamic model, has been the basis of work by Goldhaber and Teller (1948), Steinwedel, Jensen, and Jensen (1950), Danos (1952), and Araujo (1954). These theories consist essentially in variations of the original ideas of Goldhaber and Teller, in which vibrations were set up within the nucleus by the action of the radiation. The model they considered in most detail assumed that a spherical region of "proton fluid" vibrated in opposition to a similar, nearly coincident sphere of "neutron fluid", this vibration having a natural frequency corresponding to the energy of the giant resonance. Steinwedel, Jensen, and Jensen (1950) have considered the case where the target nucleus is spherical, and remains spherical even under excitation by γ -rays. Motion of the nucleons within the nucleus causes density changes inside a rigid spherical surface. The total nuclear density is assumed to be constant, while allowing local complementary periodic changes in neutron or proton density.

* Physics Department, University of Melbourne.

Using calculations for the hydrodynamics of two fluids in a sphere, the results obtained by Steinwedel, Jensen, and Jensen predict that:

$$E_m \text{ should be proportional to } A^{-\frac{1}{2}},$$

$$\int \sigma(E) dE \text{ should be proportional to } A,$$

where $\sigma(E)$ is the absorption cross section at energy E , and A is the mass number of the nucleus in question. Experimental work so far does not definitely distinguish between this predicted dependence of E_m on A and the Goldhaber-Teller prediction of an $A^{-\frac{1}{2}}$ variation.

II. THE RESONANCE WIDTH

A feature of the half-width, Γ MeV, of the giant resonance which was examined by Wilkinson (1956) is its variation from one nucleus to another. Nathans and Halpern (1953) have observed that for nuclei with closed shell structure the resonance widths are unusually small, and Wilkinson has given a qualitative account of factors which could contribute to this effect. As one of the factors, he has suggested that rare earth nuclei not having closed shells may have a broader resonance, because of a breakdown of the j - j coupling model of shell structure considered in his model. However, he has not offered a quantitative explanation.

Recently, Danos (1956, 1958) pointed out that by considering a hydrodynamic model, it is possible to predict quantitatively that the giant resonance in deformed nuclei is, in fact, the sum of two resonances at different energies. This would lead one to expect that the observed resonance in these nuclei would be broadened or double-peaked, depending on the shapes and energy separation of the two components. The resonance width for nuclei with closed or nearly closed shells (and therefore of spherical shape) would not exhibit this effect. The detailed shape of the resonance in deformed nuclei is thus of particular interest, as it provides one experimental test of the ability of the hydrodynamic model to describe the behaviour of nuclear matter. The other principal features of the giant resonance are not sensitive to the theory used to predict them. Therefore a test of any model cannot be based upon their measurement (Montalbetti, Katz, and Goldemberg 1953).

Danos (1956, 1958) has applied to the model of Steinwedel, Jensen, and Jensen (1950) boundary conditions representing a spheroidal nucleus having semi-axes a and b , where a is directed along the axis of rotational symmetry. In the solution of the eigenvalue equations in ρ_p , the proton density, to find the frequencies ω_a and ω_b of the lowest vibration modes, it is found that the eigenvalues ω depend on the orientation of the vibration considered, and also on the deformation parameter ϵ , which is defined by

$$\epsilon = (a^2 - b^2)/R^2,$$

where R is the radius of a sphere equal in volume to the spheroid. By computation over a range of ϵ -values, the following empirical formula has been shown to give accurately the splitting of the eigenvalues (Danos 1958),

$$\omega_b/\omega_a = 0.911a/b + 0.089.$$

ω_a here represents the resonance frequency for absorption of polarized γ -rays by a nucleus whose a -axis coincides with the direction of polarization, and similarly for ω_b . Curves of absorption cross section for these polarized beams would then have the usual resonance shape, peaked at ω_a and ω_b respectively, and having cross sections $\sigma_a(E)$ and $\sigma_b(E)$. For randomly oriented nuclei, the effective total cross section at energy E is expected to be

$$\langle \sigma(E) \rangle = \left\{ \frac{1}{3} \sigma_a(E) + \frac{2}{3} \sigma_b(E) \right\},$$

since there are two axes of length b and only one of length a . The total cross section should thus be the sum of two resonance curves, separated in energy, with one twice as high as the other if it is assumed that these resonance curves have the same width. More generally, the areas under these resonance curves should be related by the equation

$$\int \sigma_a(E) dE / \int \sigma_b(E) dE = \frac{1}{2}.$$

In a nucleus with ε greater than zero, the larger peak will occur at the higher energy.

In the present experiment, a careful study has been made of the absorption cross section of a nucleus having a large deformation in the ground state. Work involving experimental readings at small energy intervals is essential in order to be able to detect for certain the presence or absence of the broadening or double-peaking effect. In much other work of a similar nature, energy intervals of 1 MeV or larger have been used, and the effect could easily have been missed.

III. THE EXPERIMENT

An examination of the resonance in photon absorption by a deformed nucleus has been undertaken. In a heavy nucleus (say Z greater than 50), the sum of $\sigma(\gamma, n)$ and $\sigma(\gamma, 2n)$ gives a very good approximation to the absorption cross section, assuming that the contribution of competing reactions in which protons are emitted is small enough to be neglected. This is so for the heavy nuclei. For example, Toms and Stephens (1955) found that the yield of protons emitted from tantalum under the action of 23 MeV bremsstrahlung was less than 10^{-3} of the corresponding neutron yield.

Since the intrinsic quadrupole moment of a nucleus (Q_0) is directly proportional to ε , an element with a high quadrupole moment was used as target. In this case tantalum 181, a 100 per cent. natural isotope, was chosen. (Further tests are to be done on some rare earths, which are equally suitable.) Since the nucleus ^{181}Ta has Z (=73) and N (=108) well away from the magic number values of 50, 82, and 126, one would expect this nucleus to be deformed.

The spectroscopically measured electric quadrupole moment Q_s is related to the intrinsic moment Q_0 by the equation

$$Q_0 = Q_s \frac{(I+1)(2I+3)}{I(2I-1)},$$

where I is the nuclear spin. Q_s has been determined for ^{181}Ta by Murakawa and Kamei (1957), who find Q_s to be $+2.7 \pm 0.3$ barns. Since for tantalum, I is $7/2$,

Q_0 is calculated to be $+5.8$ barns. A determination of Q_0 from probabilities of γ -ray transitions found in Coulomb excitation of low-lying states has given a value of $+6.8$ barns (Alder *et al.* 1956).

Since in the ^{181}Ta nucleus Q_0 is positive, a is greater than b . Thus, if the predictions of Danos are correct, the absorption cross section resonance should be broadened or double-peaked, due to the two components, and the higher energy component should have the larger $\int \sigma(E)dE$. For example, taking the low energy peak to be at 13 MeV, and Q_0 to be $+6.0$ barns, the theory indicated a separation in peaks of about 2.8 MeV.

Bremsstrahlung from the Melbourne 18 MeV electron synchrotron was used for the irradiation of the tantalum target. Maximum X-ray energy was variable between 8 and 18 MeV by means of variation and precise measurement of the d.c. bias used in exciting the magnet. An energy calibration of the machine by means of (γ, n) threshold measurements on ^{31}P , ^{32}S , ^{55}Mn , ^{63}Cu , and ^{65}Cu had been completed one month prior to this work, and the effective X-ray energies were considered to be known absolutely within ± 0.2 MeV, with relative energies defined considerably better. The X-ray pulse was obtained in each cycle by cutting off the r.f. accelerating field 100 μsec before the peak of the a.c. cycle. The electron beam would then spiral inwards onto a platinum target 0.005 in. thick. The length of time (250 μsec) over which the bulk of the beam emerged in each cycle has allowed gating circuits to be omitted from the counting electronics, there being no significant pile-up of electron pulses in the counters.

The arrangement of target, neutron counters, and paraffin and concrete "house" was similar to that described by Nathans, Halpern, and Mann (1952). The target consisted of a block of better than 99.9 per cent. pure tantalum, 1 in. square and 0.14 in. long in the direction of the beam. In the analysis of results, correction was made for the absorption of X-rays of different energies over the length of this block.

A pair of ^{10}B -enriched boron trifluoride proportional counters was used to count neutrons thermalized by the 13 cm of paraffin separating the counters from the target. At this distance, according to Rossi and Staub (1949), the counting efficiency will be energy independent. The counter pulses were fed through an amplifying and discriminating system to a pair of independent scalars. Counting efficiency was checked regularly throughout the work by measuring the count rate from a 10 mc Ra-Be neutron source which could be inserted in the paraffin behind the target. Variations in counting efficiency were not greater than ± 1 per cent. from the mean value over a day's running.

The dead-time inherent in the system was the source of a counting loss which depended on the average input pulse rate. To take this into consideration reliably, a correction factor was determined experimentally as a function of average counting rate, and was applied to every point of the yield curve.

The radiation dose for each exposure was measured with a Victoreen 0.25 r integrating monitor, used as a transmission chamber and placed behind the lead collimator so as to intercept the whole of the collimated beam. This ionization chamber was tested periodically for leakage, and was calibrated in terms of

"Lucite roentgens" by a comparison of its response with that of a 25 r thimble irradiated at the centre of an 8 cm "Perspex" cube.

Synchrotron energy was varied in steps of 0.2 MeV over the range 8–18 MeV to obtain a yield curve in terms of neutron counts per unit dose versus synchrotron energy. Except close to the $\text{Ta}(\gamma, n)$ threshold, between 4000 and 10,000 counts were taken per run. Each run was repeated independently not less than four times, and average yields were plotted. Thus the statistical

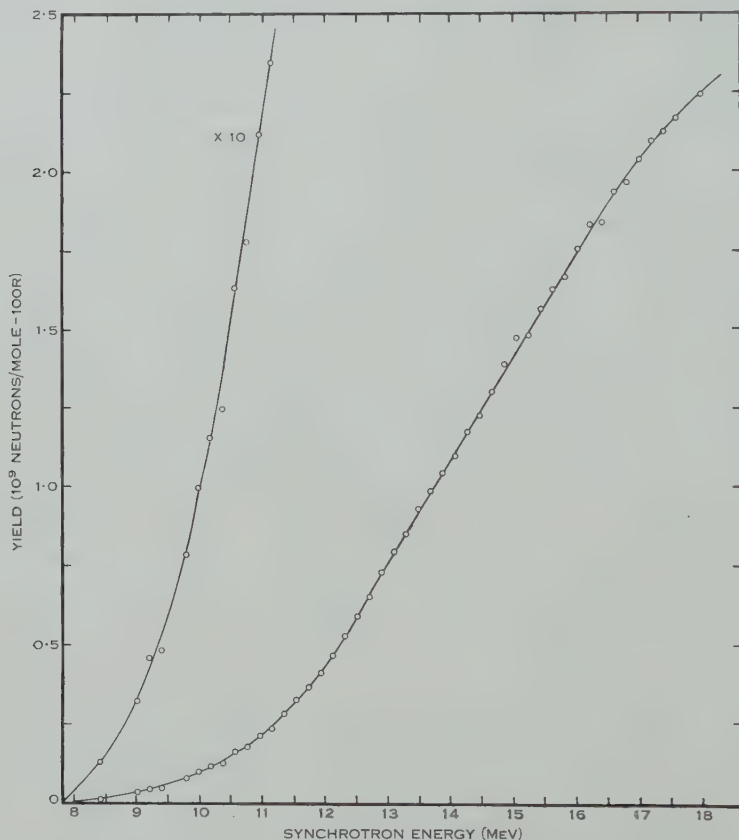


Fig. 1.—Measured yield curve of photoneutrons from tantalum as a function of synchrotron energy.

counting error was between 0.5 and 1 per cent., except near the threshold. From this graph was subtracted a similarly obtained background curve, measured with all apparatus in place except the tantalum itself. At 17 MeV, it represented 4 per cent. of the total neutron yield.

A plot of the low energy yield curve points has indicated a value for the $\text{Ta}(\gamma, n)$ threshold of 7.7 ± 0.2 MeV. This agrees with figures previously obtained by McElhinney *et al.* (1949) (7.7 ± 0.2 MeV), by Sher, Halpern, and Mann (1951) (7.55 ± 0.20 MeV), and by Cameron (1957), whose mass formula calculation gave 7.77 MeV.

The absolute yield curve, obtained by taking account of counter efficiency and absorption of X-rays in the target, is shown in Figure 1. Since the test neutron source calibration was quoted to ± 10 per cent., the absolute values may be in error by this amount. However, the relative values will be very much better than this, because of the good counting statistics.

IV. ANALYSIS OF RESULTS

Derivation of the cross section from the yield curve was carried out by the Leiss-Penfold method (Penfold and Leiss 1954), using effective energy intervals of 0.5 MeV (Fig. 2). First differences of the yield curve were smoothed, to minimize errors due to the drawing of a curve through the experimental points.

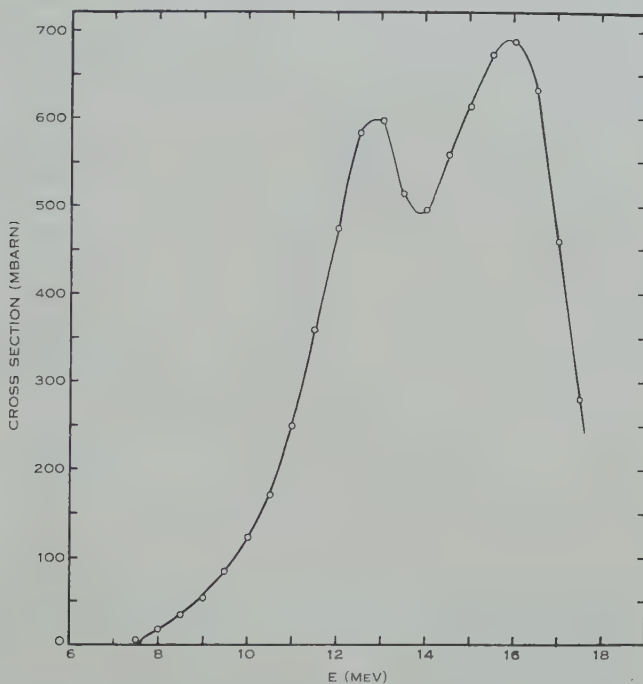


Fig. 2.—Cross section calculated by the Leiss-Penfold method from the neutron yield curve. This cross section plot represents $\sigma(\gamma, n) + 2\sigma(\gamma, 2n)$ as a function of energy.

Above the threshold of the $^{181}\text{Ta}(\gamma, 2n)$ reaction, account has to be taken of multiple neutron production, since the cross section derived from the neutron yield curve is the sum of $\sigma(\gamma, n) + 2\sigma(\gamma, 2n) + 3\sigma(\gamma, 3n)$ and so on. To obtain the absorption cross section, the extra weighting of multiple neutron reactions must be allowed for. For our purposes, the absorption cross section is $\sigma(\gamma, n) + \sigma(\gamma, 2n)$, since the threshold for the $^{181}\text{Ta}(\gamma, 3n)$ reaction is greater than 18 MeV.

There being very little experimental evidence available on the subject, statistical theory of nuclear reactions as described by Blatt and Weisskopf (1952) was used. The result which they derive is

$$\sigma(\gamma, 2n) = \sigma_{\text{abs}} \{1 - (1 + \epsilon_{\text{sec}}/\Theta) \exp(-\epsilon_{\text{sec}}/\Theta)\},$$

where σ_{abs} is the total cross section for nuclear absorption of a γ -ray. ϵ_{sec} is the maximum energy which could be given to a second neutron emitted from the nucleus, that is, it is equal to the difference between the γ -ray energy and the energy of the $(\gamma, 2n)$ threshold. Θ is the maximum possible nuclear temperature of the ^{180}Ta nucleus, which is formed after the emission of the first neutron, and before the second neutron is emitted. Θ is defined by the equation

$$\Theta = \frac{1}{a} \{E_\gamma - E_{\text{th}}\}^{\frac{1}{2}},$$

where a is a constant, estimated by Blatt and Weisskopf (1952) to have the value 10 in our case. E_{th} is the (γ, n) threshold energy.

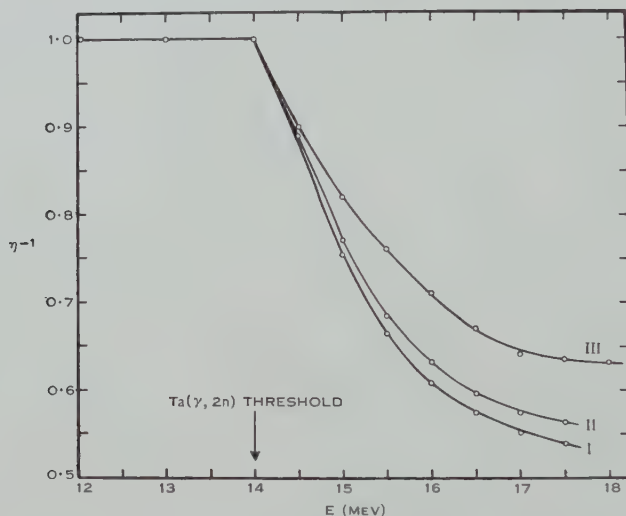


Fig. 3.—The reciprocal of the neutron multiplicity plotted as a function of γ -ray energy for tantalum. Curve I shows $(\eta_1)^{-1}$ calculated on the basis of the statistical theory of nuclear reactions. Curve II shows $(\eta_2)^{-1}$ calculated on statistical theory plus an assumed 10 per cent. of the absorption giving rise to direct photoeffect. Curve III shows $(\eta_3)^{-1}$ as found in the experiment of Whalin and Hanson (1953).

This formula for $\sigma(\gamma, 2n)$ is based on the assumptions that $E_\gamma \gg \Theta$, and that neutron emission from the intermediate nucleus predominates immediately it becomes energetically possible. Both of these assumptions are good for our case.

In our calculation, a was put equal to 10, and the values of the (γ, n) and $(\gamma, 2n)$ thresholds used were 7.7 and 14.0 MeV respectively. The result of the calculation was expressed in terms of a factor $\eta(E)$, known as the neutron multiplicity. This is defined as the average number of neutrons emitted for each photon absorbed. The measured neutron yield is then given by

$$Y(E_0) = 0.6023 \int_0^{E_0} \sigma(E) \eta(E) P(E, E_0) dE,$$

where $P(E, E_0)dE$ is the number of photons of energy E to $E+dE$ per cm^2 per 100 r in a spectrum of maximum energy E_0 . The cross section is in barns in this formula. The neutron multiplicity can be less than unity owing to competition from proton emission, but, since the proton yield from tantalum is so small relative to the neutron emission, the minimum value of $\eta(E)$ in our case is unity. A graph of $1/\eta(E)$ for tantalum as a function of γ -ray energy is shown in Figure 3. The calculation using statistical theory leads to the curve $(\eta_1)^{-1}$ in that figure. Thus the absorption cross section is η^{-1} times the cross section derived from the neutron yield curve.

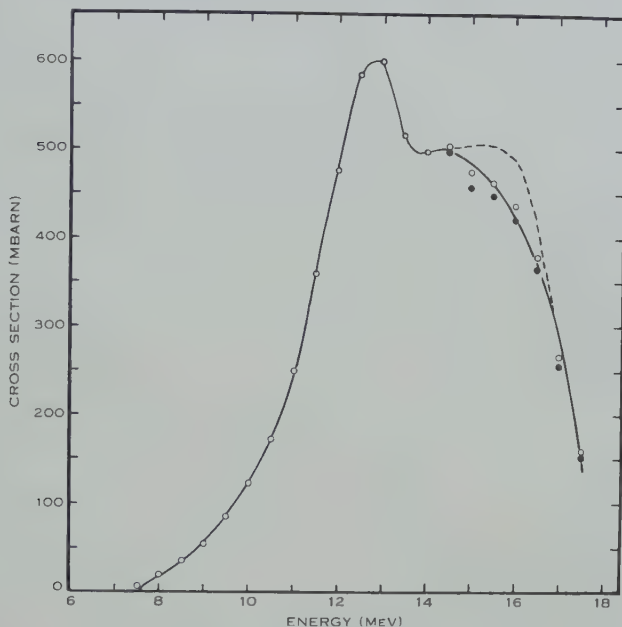


Fig. 4.—The cross section for nuclear absorption of γ -rays in tantalum. The full circles were calculated using curve I of Figure 3, while the open circles were the result of using curve II. Use of curve III led to the dotted cross section.

An attempt to allow for the effect of the direct photoeffect was made. Once again, experimental evidence was lacking, and to estimate the significance of the effect the cross section for direct interaction between photons and nucleons of the tantalum nucleus was assumed to be a constant 10 per cent. of the total cross section for absorption of photons by that nucleus. Including this consideration, an alternative factor was found, and this is plotted as $(\eta_2)^{-1}$ in Figure 3.

Experimental values for the ratio of $\sigma(\gamma, n)$ to $\sigma(\gamma, 2n)$ have been obtained by Whalin and Hanson (1953) and by Carver, Edge, and Lokan (1957). This was done by measuring a total neutron yield curve, and subtracting from it the yield curve for the (γ, n) reaction only. The (γ, n) yield was obtained by counting the β^+ -activity of ^{180}Ta . This subtraction gives the yield curve for the $(\gamma, 2n)$ reaction. These results disagree with the statistical theory very considerably.

The values of η^{-1} obtained in the experiment of Whalin and Hanson are plotted in Figure 3 as $(\eta_3)^{-1}$.

The effects of η_1 and η_2 respectively are shown in the two sets of plotted points in Figure 4, which shows the cross section for photon absorption in the tantalum nucleus. Precisely where the actual curve should lie cannot be said. However, the essential features of the final cross section are not at all dependent on these doubtful multiplying factors. If the experimental values of $\eta(E)$ are

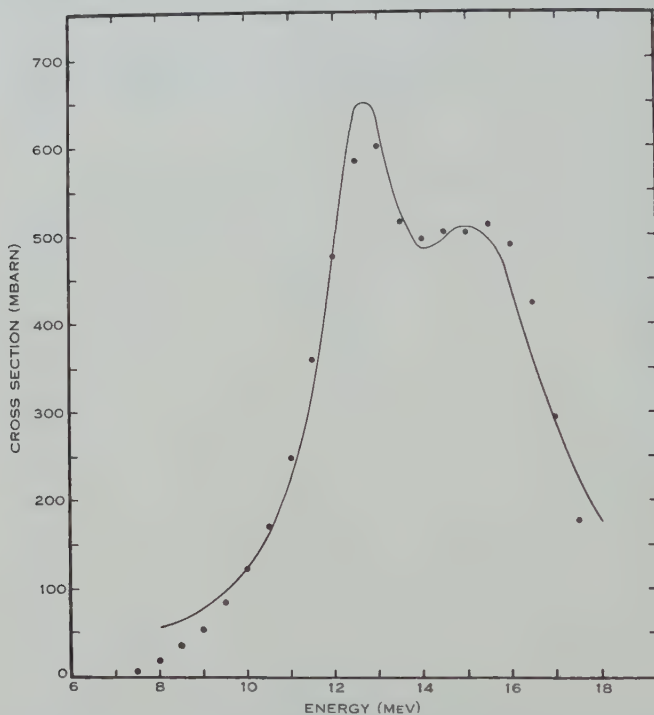


Fig. 5.—The fit of two Breit-Wigner shape resonance curves to the cross section calculated using curve III of Figure 3. Parameters of Breit-Wigner fit to observed σ : $E_1=12.6$ MeV, $\sigma_{1m}=500$ mbarn, $\Gamma_1=2$ MeV; $E_2=15.3$ MeV, $\sigma_{2m}=450$ mbarn, $\Gamma_2=4$ MeV.

used, the dotted curve of Figure 4 is obtained. In this case, the evidence for a second peak in the cross section for the absorption of γ -rays is even more pronounced.

In view of the limits of error shown and the constancy of the characteristic general shape of the cross section under different methods of allowing for the $(\gamma, 2n)$ reaction, it is considered that the evidence for two peaks in the absorption cross section is quite definite. The width of the resonance curve of Figure 4 (full width at half maximum) is 5.4 MeV, the maximum cross section is 615 mbarn, and the integrated cross section up to 18 MeV is 3.2 MeV-barn.

There remains some doubt about our assessment of the absolute yield, mainly owing to the uncertainty in calibration of the neutron source used to test the counter efficiency and to uncertainty in the calibration of the Victoreen

thimble as used in the 10–20 MeV region. If instead we standardize our neutron yield from tantalum against that from copper, and use the value quoted by Montalbetti, Katz, and Goldemberg (1953) for the neutron yield from this element, the peak cross section becomes 540 mbarn.

V. CONCLUSIONS

The γ -ray absorption resonance observed for tantalum exhibits distinct features of shape and width not normally found in curves of this type. This result can be satisfactorily explained in terms of a sum of two components, and a "resonance" showing just these features might well be expected on the basis of the theory of Danos (1958). To demonstrate the feasibility of such a resolution

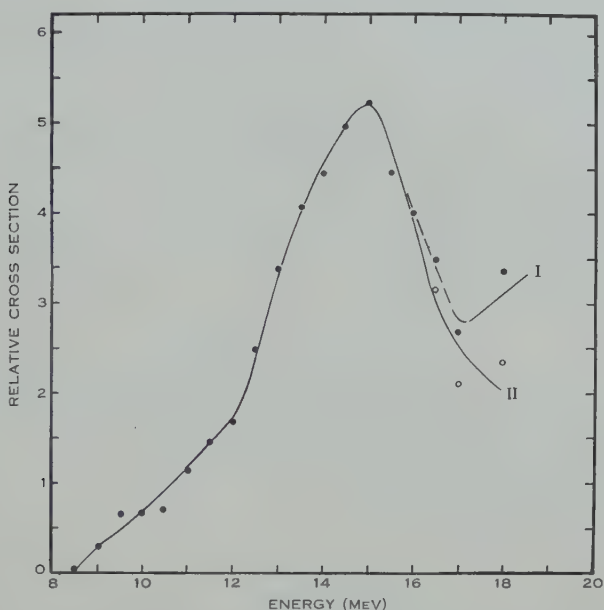


Fig. 6.—Cross section for nuclear absorption of γ -rays in tantalum. Curve I is the cross section calculated from the neutron yield curve. Curve II is corrected for the effects of neutron multiplicity.

into two peaks as required by the Danos theory, the cross section obtained using the experimental values of $\eta(E)$ has been decomposed into two components, σ_1 and σ_2 (see Fig. 5). Both these component cross sections were assumed to have Breit-Wigner shape, with parameters as stated in Figure 5. This decomposition is not unique, of course, but serves to show the compatibility of this result with the predictions of Danos (1958). The energy separation of the two peaks in this decomposition leads to a value of the intrinsic electric quadrupole moment for tantalum of 6.1 barns (compared with the previously quoted values of 5.8 and 6.8 barns). This value was obtained assuming $R=R_0A^{\frac{1}{3}}$, $R_0=1.2 \times 10^{-13}$ cm. Further, the ratio of area under the lower energy peak

to that under the higher energy peak is 0.55, in approximate agreement with the predicted value of 0.5.

Whilst the above results thus lend clear support to the validity of Danos's improvement to the hydrodynamic model, it is not claimed that this theory is the correct or the only interpretation of them. Before any adequate verification of the theory is obtained, the γ -ray absorption cross sections must be measured for a number of nuclei of varying deformations. In particular, it must be demonstrated that this effect does not occur for undeformed nuclei. To test this, the above technique has been applied to the measurement of the γ -ray absorption cross section in lanthanum. This is a nucleus with a closed shell of neutrons ($N=82$) and thus has very small deformation. The absorption cross section obtained for this nucleus is shown in Figure 6, and shows no evidence whatever for two components. This is also in agreement with the hydrodynamic model theory.

We are grateful to have been informed of a series of experiments, with an object identical to the present one, recently performed by Weiss and Fuller (Fuller, personal communication 1957). The cross section for absorption of γ -rays by the tantalum nucleus has been resolved into two peaks at 12.8 and 15.7 MeV respectively. The integrated cross section to 20 MeV was found to be 3.27 MeV-barn and the width of the resonance about 6 MeV. This is in good agreement with the present work.

VI. ACKNOWLEDGMENTS

The authors wish to express their thanks to Professor Sir Leslie Martin for his interest and encouragement during this work. Two of us (H.H.T. and F.R.A.) wish to thank Imperial Chemical Industries of Australia and New Zealand Ltd. and Dunlop Rubber Australia Ltd. respectively for research scholarships held during this work.

VII. REFERENCES

- ALDER, K., BOHR, A., HUUS, T., MOTTELSON, B. R., and WINTHER, A. (1956).—*Rev. Mod. Phys.* **28**: 432.
- ARAUJO, J. M. (1954).—*Nuovo Cimento* **12**: 780.
- BLATT, J. M., and WEISSKOPF, V. F. (1952).—"Theoretical Nuclear Physics." p. 379. (Wiley: New York.)
- BURKHARDT, J. L. (1953).—*Phys. Rev.* **91**: 420.
- CAMERON, A. G. W. (1957).—Atomic Energy of Canada Ltd. Report CRP 690.
- CARVER, J. H., EDGE, R. D., and LOKAN, K. H. (1957).—*Proc. Phys. Soc. Lond. A* **70**: 415.
- DANOS, M. (1952).—*Ann. Phys., Lpz.* **6**: 265.
- DANOS, M. (1956).—*Bull. Amer. Phys. Soc.* **1**: 135.
- DANOS, M. (1958).—*Nuclear Physics* **5**: 23.
- GOLDHABER, M., and TELLER, E. (1948).—*Phys. Rev.* **74**: 1046.
- LEVINGER, J. S., and BETHE, H. A. (1950).—*Phys. Rev.* **78**: 115.
- LEVINGER, J. S., and KENT, D. C. (1954).—*Phys. Rev.* **95**: 418.
- McELHINNEY, J., BECKER, R. A., HANSON, A. O., DUFFIELD, R. B., and DIVEN, B. C. (1949).—*Phys. Rev.* **75**: 542.
- MONTALBETTI, R., KATZ, L., and GOLDBERG, J. (1953).—*Phys. Rev.* **91**: 659.
- MURAKAWA, K., and KAMEI, T. (1957).—*Phys. Rev.* **105**: 671.
- NATHANS, R., and HALPERN, J. (1953).—*Phys. Rev.* **92**: 207.

- NATHANS, R., HALPERN, J., and MANN, A. K. (1952).—*Rev. Sci. Instrum.* **23**: 678.
- PENFOLD, A. S., and LEISS, J. F. (1954).—*Phys. Rev.* **95**: 637.
- ROSSI, B., and STAUB, H. (1949).—"Ionisation Chambers and Counters." p. 192. (McGraw-Hill: New York.)
- SHER, R., HALPERN, J., and MANN, A. K. (1951).—*Phys. Rev.* **84**: 387.
- STEINWEDEL, H., JENSEN, J. H. D., and JENSEN, P. (1950).—*Z. Naturf.* **5a**: 343, 413.
- TOMS, M. E., and STEPHENS, W. E. (1955).—*Phys. Rev.* **98**: 626.
- WHALIN, E. A., and HANSON, A. O. (1953).—*Phys. Rev.* **89**: 324.
- WILKINSON, D. H. (1956).—*Physica* **22**: 1039.

THE FLUCTUATING FIELD FERROMAGNET AT LOW TEMPERATURES

By F. D. STACEY*

[Manuscript received February 27, 1958]

Summary

Néel's concept of fluctuating intermolecular field may be ascribed to localized but migratory elementary magnets in ferromagnetics with non-integral Bohr magneton numbers. This concept allows a low temperature ordering of elementary magnets and a variation with temperature of the effective number of nearest neighbours between which exchange interactions occur. This model alone cannot explain the approach to absolute saturation at 0 °K of a ferromagnetic, but it modifies the constant of the $T^{3/2}$ law of approach. In the case of ordering by the mutual attraction of parallel elementary magnets, each consisting of a coupled pair of spins ($s=1$) the theoretical and experimental values of this constant for nickel are reconciled.

I. INTRODUCTION

The statistical nature of the alignment of elementary magnets in a ferromagnetic domain cannot be properly accounted for in any existing theory. Néel (1932, 1934, 1940) directed attention to this deficiency of the theory of ferromagnetism by considering statistical fluctuations of intermolecular field near to the Curie temperature. He obtained a convincing correlation between specific heat, the difference between paramagnetic, θ_p , and ferromagnetic, θ_f , Curie temperatures, and the departure from linearity of the $(1/\chi - T)$ curve above θ_f . The fluctuations are considered to arise by thermal migration of the elementary magnets from one lattice site to another, being discretely located on particular sites at any instant.

In an earlier paper (Stacey 1955) this model was examined for a face-centred cubic lattice to determine the conditions under which ordering of elementary magnets on the lattice sites could explain the observed variation of intermolecular field in nickel. An analytical form for the approach to absolute saturation at 0 °K did not appear to be obtainable for this lattice, so that an important deficiency of the model was missed. The body-centred cubic lattice contains only two (instead of four) sublattices and is therefore algebraically easier to handle. It gives an exponential approach to absolute saturation, clearly in disagreement with experimental results for nickel. (A qualitative similarity between the behaviour of face-centred and body-centred lattices must be expected.)

The success of Bloch's (1932) law, $1 - \sigma/\sigma_0 = CT^{3/2} + \dots$, for the departure of spontaneous magnetization σ/σ_0 from saturation at low temperatures, T , is explained by Dyson's (1956*b*) paper in which the next term is shown to be proportional to $T^{5/2}$ and not a lower power as had been reported earlier. This

* Geophysics Department, Australian National University, Canberra.

allows the $T^{3/2}$ law quite a wide range of validity as observed experimentally, but the theoretical and experimental values of C do not agree. C depends directly upon the number, η , of nearest neighbouring lattice sites between which exchange interactions occur, and the assumption is made that η has the same value at 0 °K and the Curie point. An estimate is therefore readily made of the correction to C indicated by the fluctuating field concept.

II. LATTICE ORDERING BY MUTUAL REPULSION OF ELEMENTARY MAGNETS

In this section consideration is given to the behaviour at low temperatures of a body-centred cubic lattice, incompletely occupied by elementary magnets, between which there is an ionic repulsion exceeding their exchange interaction.

The lattice can be divided into two completely equivalent sublattices, which, in any unit cell, are represented by the cube corners and the body centre. Consider two sublattices with different populations w and x of elementary magnets, disposed on a total of N lattice sites, where

$$w+x=n, \quad \dots\dots\dots (1)$$

n being the total number of magnets, and for convenience

$$w-x=n\delta. \quad \dots\dots\dots (2)$$

The w and x magnets will have different spontaneous magnetizations $(\sigma/\sigma_0)_w$ and $(\sigma/\sigma_0)_x$, and we may put

$$W=w(\sigma/\sigma_0)_w, \quad X=x(\sigma/\sigma_0)_x,$$

so that

$$W+X=n\sigma/\sigma_0, \quad \dots\dots\dots (3)$$

$$W-X=n(\sigma/\sigma_0)\Delta. \quad \dots\dots\dots (4)$$

Each x site has eight nearest neighbours, all on the w sublattice, and similarly each w site has eight nearest neighbouring x sites. Therefore, if next nearest neighbour interactions are neglected, the intermolecular field acting on each sublattice is due to the occupation and spontaneous magnetization of the other sublattice. Equations representing the interdependence of the spontaneous magnetizations are derived in the same way as for the simple quantum-corrected Weiss theory. Equations are numbered (5A), (6A), etc. for the case of elementary magnets being single spins ($s=\frac{1}{2}$) and (5B), (6B), etc. for coupled spins ($s=1$).

$$\frac{W}{w} = \tanh \left(\frac{16J}{NkT} X \right), \quad \dots\dots\dots (5A)$$

$$\frac{X}{x} = \tanh \left(\frac{16J}{NkT} W \right), \quad \dots\dots\dots (6A)$$

$$\frac{2 - \sqrt{4 - 3(W/w)^2}}{W/w} = \tanh \left(\frac{8J}{NkT} X \right), \quad \dots\dots\dots (5B)$$

$$\frac{2 - \sqrt{4 - 3(X/x)^2}}{X/x} = \tanh \left(\frac{8J}{NkT} W \right), \quad \dots\dots\dots (6B)$$

where J represents the exchange energy between occupied neighbouring sites, T is temperature, and k Boltzmann's constant. Equations (5) and (6) refer to spontaneous magnetization in zero external field.

Both lattices have the same Curie temperature, θ , since both W and X are positive and by (5) and (6) they must approach zero together. Simplification of these equations for $W, X \rightarrow 0$ gives

$$\theta = \frac{8J}{k} \frac{n}{N}, \quad \dots\dots\dots (7A)$$

$$\theta = \frac{2}{3} \cdot \frac{8J}{k} \frac{n}{N}. \quad \dots\dots\dots (7B)$$

For approximation to equations (5) and (6) at low temperatures, the expansion

$$\tanh A = 1 - 2e^{-2A} + \dots$$

is used. Substituting equations (1)–(4) and (7) in (5) and (6) we obtain at low temperatures :

$$1 - \frac{\sigma}{\sigma_0} = (1 + \delta) \exp \left\{ -2 \frac{\theta}{T} \frac{\sigma}{\sigma_0} (1 - \Delta) \right\} + (1 - \delta) \exp \left\{ -2 \frac{\theta}{T} \frac{\sigma}{\sigma_0} (1 + \Delta) \right\}, \quad \dots\dots\dots (8A)$$

$$1 - \frac{\sigma}{\sigma_0} = \frac{1}{2} \left(1 + \delta - \frac{\sigma}{\sigma_0} \cdot \frac{1 + \Delta}{2} \right) \exp \left\{ -\frac{3}{2} \frac{\theta}{T} \frac{\sigma}{\sigma_0} (1 - \Delta) \right\} \\ + \frac{1}{2} \left(1 - \delta - \frac{\sigma}{\sigma_0} \cdot \frac{1 - \Delta}{2} \right) \exp \left\{ -\frac{3}{2} \frac{\theta}{T} \frac{\sigma}{\sigma_0} (1 + \Delta) \right\}, \quad \dots\dots\dots (8B)$$

provided that δ and Δ do not approach unity, i.e. n/N must be greater than $\frac{1}{2}$.

The behaviour of δ and Δ as $T/\theta \rightarrow 0$ is readily determined by considering the free energy, F , of the system.

$$F = U - kT \ln A,$$

where U is internal energy and A is the number of complexions of the system. When the spontaneous magnetizations of both sublattices approximate to unity, U and A are given by

$$U = \frac{16wx}{N} (E - J), \quad \dots\dots\dots (9) \\ A = \frac{(\frac{1}{2}N)!}{(\frac{1}{2}N - w)! w!} \cdot \frac{(\frac{1}{2}N)!}{(\frac{1}{2}N - x)! x!},$$

so that

$$\frac{1}{kT} \frac{\partial F}{\partial x} = \frac{16(E - J)}{NkT} (w - x) + \ln \left(\frac{x(\frac{1}{2}N - w)}{w(\frac{1}{2}N - x)} \right),$$

which is equated to zero for the required minimum of free energy.

This causes w and x to approach limiting values $\frac{1}{2}N$ and $(n-\frac{1}{2}N)$ in an exponential manner as $T \rightarrow 0^\circ\text{K}$, so that at low temperatures we may write $\delta = \Delta = (N/n - 1)$, whence

$$1 - \frac{\sigma}{\sigma_0} = \frac{N}{n} \exp \left\{ -2 \left(2 - \frac{N}{n} \right) \frac{\theta}{T} \right\} + \left(2 - \frac{N}{n} \right) \exp \left\{ -2 \frac{N}{n} \frac{\theta}{T} \right\}, \quad \dots\dots\dots (10A)$$

$$1 - \frac{\sigma}{\sigma_0} = \frac{1}{4} \frac{N}{n} \exp \left\{ -\frac{3}{2} \left(2 - \frac{N}{n} \right) \frac{\theta}{T} \right\} + \frac{1}{4} \left(2 - \frac{N}{n} \right) \exp \left\{ -\frac{3}{2} \frac{N}{n} \frac{\theta}{T} \right\}. \quad \dots\dots\dots (10B)$$

In equations (10) the first terms predominate. Subsequent terms in the expansions involve squares of the exponentials which are not significant below $T/\theta = 0.2$. Over this temperature range the $T^{3/2}$ law of Bloch (1932) gives a much better fit to experimental values than any combination of exponentials in the form of equation (9), which shows the unsuitability of this approach at low temperatures, and suggests an examination of the $T^{3/2}$ law.

III. THE $T^{3/2}$ LAW

Dyson (1956*a*, 1956*b*) gave a complete analysis of the spin-wave method of calculating the spontaneous magnetization, σ/σ_0 , of a ferromagnetic at low temperatures. The important conclusion from these calculations is that the next term in the expansion of σ/σ_0 in powers of temperature, T , is proportional to $T^{5/2}$. Thus there is a useful range of temperatures over which the $T^{3/2}$ term expresses σ/σ_0 versus T as accurately as experiment can check it. Available experimental data do not allow a very close check of the $T^{3/2}$ law but Arrott (1955) has provisionally reported that more accurate measurements are being made.

Combining equations (135) and (143) in the paper by Dyson (1956*b*) we have at low temperatures

$$1 - \frac{\sigma}{\sigma_0} = \frac{0.1659}{s} \left(\frac{T}{T_c} \right)^{3/2} + \dots, \quad \dots\dots\dots (11)$$

where $2s$ is the number of spins per elementary magnet, and T_c is a constant approximately equal to the Curie temperature θ . To give reasonable agreement with experiment we require $s = 3/2$ in nickel, which is generally regarded as improbable. An alternative explanation for the value of the constant in equation (11) is found in the fluctuating field theory.

The quantity T_c in equation (11) is proportional to γ_0 , the number of nearest neighbours per atom in Dyson's theory, but what appears to be significant is the average number, η , of occupied nearest neighbours to each of the lattice sites occupied by an elementary magnet. If there are fewer elementary magnets than lattice sites, η may be temperature dependent, which γ_0 is not.

The range of variation of η with temperature was calculated for nickel with $s = \frac{1}{2}$ and $s = 1$ by Stacey (1955), assuming repulsive ordering of elementary magnets on a face-centred cubic lattice. It must be noted that the values for

the ordered and disordered states were accidentally inverted and that η is smaller in the low temperature ordered state. Without numerical calculation it can be seen immediately that this process would increase the disagreement between the theoretical and experimental constants of equation (11). We have

$$T_c = f \frac{\eta_0}{\gamma_0} \theta, \dots\dots\dots (12)$$

where η_0 and γ_0 are the values of η at $T=0$ and $T=\theta$, and f is a numerical factor approximately equal to unity.

For equation (11) to agree with the experimental equation

$$1 - \sigma/\sigma_0 = 0.11(T/\theta)^{3/2}, \dots\dots\dots (13)$$

we require $\eta_0/\gamma_0 > 1$, which can only result from what is here termed "attractive ordering".

IV. LATTICE ORDERING BY MUTUAL ATTRACTION OF ELEMENTARY MAGNETS

The type of order which follows from the tendency of elementary magnets to choose as neighbours magnets parallel to themselves is here examined for a face-centred cubic lattice, and η_0/γ_0 is calculated for $s=\frac{1}{2}$ and $s=1$ in nickel. Attractive ordering instead of repulsive ordering is to be expected if $(E-J)$ (equation (9)) is negative.

Consider a face-centred cubic lattice of N sites, arranged in a cube of side $(\frac{1}{4}N)^{1/3}.a$, where a is the side of a unit cell and N is large. There are n elementary magnets disposed on the N sites. Since each site has 12 nearest neighbours, in a disordered arrangement of n magnets the average number of nearest neighbours is

$$\eta_0 = 12n/N. \dots\dots\dots (14)$$

This arrangement is approximately valid at the Curie point, θ .

At low temperatures the magnets will tend to become ordered into co-planar groups, the planes being parallel to the faces of the unit cell. This arrangement increases the number of nearest neighbour interactions, with the minimum of local concentration of magnets. In the ordered limit the elementary magnets will occupy p complete planes in each of the three mutually perpendicular directions. The whole lattice has $2(\frac{1}{4}N)^{1/3}$ planes in each direction, each plane having $2(\frac{1}{4}N)^{2/3}$ sites. For $n/N < 0.25$ the planes would be expected to be incomplete and the following calculation would be invalid, but the interesting values are $n/N = 0.604, 0.302$, corresponding to $s = \frac{1}{2}, 1$ in nickel.

The intersection of three planes may occur either at a lattice site or between sites, depending upon the planes selected, but the first alternative is favoured for occupation as it gives a larger value of η and hence lower energy for any particular value of n/N .

With p planes in each direction there are $3p^2$ lines of intersection of two planes and p^3 intersections of three planes. Each line of intersection has $(\frac{1}{4}N)^{1/3}$ sites common to both planes which must be subtracted from the total count of

occupied sites to avoid counting twice. This would subtract twice the points of triple intersection, so that

$$n = 3p \cdot 2(\frac{1}{4}N)^{2/3} - 3p^2(\frac{1}{4}N)^{1/3} + p^3,$$

which gives

$$\frac{n}{N} = \frac{3}{2} \left[\frac{P}{(\frac{1}{4}N)^{1/3}} \right] - \frac{3}{4} \left[\frac{P}{(\frac{1}{4}N)^{1/3}} \right]^2 + \frac{1}{4} \left[\frac{P}{(\frac{1}{4}N)^{1/3}} \right]^3 \cdot \dots \quad (15)$$

It may be noticed that $n/N=1$ when half the planes are occupied, i.e. $p=(\frac{1}{4}N)^{1/3}$, alternate planes having been selected for occupation by the triple intersection condition.

Along a line of intersection between two occupied planes there are $4(\frac{1}{4}N)^{1/3}$ interactions between sites in different planes in addition to the two interactions per site within the planes. This gives a total number of interactions between nearest neighbours :

$$\frac{1}{2}n\eta_0 = 2 \cdot 3p \cdot 2(\frac{1}{4}N)^{2/3} + 3p^2 \cdot 4(\frac{1}{4}N)^{1/3},$$

$$\eta_0 = 6 \frac{N}{n} \cdot \frac{P}{(\frac{1}{4}N)^{1/3}} \left(1 + \frac{P}{(\frac{1}{4}N)^{1/3}} \right) \cdot \dots \quad (16)$$

Equations (15) and (16) can be used to find η_0 at any value of n/N . Values are given below for the cases of elementary magnets in nickel being single or coupled spins :

	$\frac{n}{N} = 0.604 \ (s=\frac{1}{2})$	$\frac{n}{N} = 0.302 \ (s=1)$
η_0 (ordered)	7.67	5.45
η_0 (disordered)	7.25	3.62

The values of η_0 are given by equation (14).

For $s=\frac{1}{2}$ the ratio η_0/η_0 is so near to unity that it cannot significantly affect the $T^{3/2}$ law, in which the constant is three times too large. For $s=1$, however, it gives an interesting agreement with the experimental law.

Combining equations (11) and (12) and using the above values of η_0 and η_0 for $s=1$ in nickel, we obtain

$$1 - \frac{\sigma}{\sigma_0} = \frac{0.08979}{f^{3/2}} \left(\frac{T}{\theta} \right)^{3/2} \cdot \dots \quad (17)$$

Since substitution of a theoretical value of f would be somewhat arbitrary, the value required for agreement between theory and experiment is obtained by comparison of equations (13) and (17) :

$$f = 0.87.$$

Dyson (1956*b*) suggests $f \sim 0.9$ as a conclusion of the calculation by Rushbrooke and Wood (1955) who give $f=0.8$ to 0.9 , so that the above value is as close as the uncertainty will permit.

V. CONCLUSION

The consideration of ionic ordering in a body-centred cubic lattice clearly indicates that the observed temperature dependence of spontaneous magnetization of a ferromagnetic near to complete saturation is not explicable in terms of variation of intermolecular field caused by the ordering. This conclusion may reasonably be applied to nickel, although the appropriate calculation would not so easily be carried out for a face-centred cubic lattice.

Dyson's (1956*a*, 1956*b*) removal of the theoretical doubt about the validity of the $T^{3/2}$ law, allows a comparison of the theoretical and experimental values of the constant C , in the equation

$$1 - \sigma/\sigma_0 = C(T/\theta)^{3/2}.$$

In nickel the agreement is not good unless one allows the elementary magnets to be spins coupled in threes. This appears improbable, although there is substantial evidence (particularly paramagnetic susceptibility) that the spins are coupled in pairs.

To explain the discrepancy on the basis of different arrangements on the lattice sites of the elementary magnets, it is necessary that a low temperature, ordered state be produced by a mutual attraction of parallel magnets. This attraction must overcome the ionic repulsion, which results from the localization of charge on atoms with vacancies in the $3d$ shells, and is contrary to an earlier calculation based on the assumption of ordering due principally to repulsion. Solid local accumulations of magnets may still be prevented by the repulsion, leading to a tendency for occupied lattice sites to form intersecting planes. This leads to an increase in the exchange energy of the lattice, and modifies the constant C . For $s = \frac{1}{2}$ the modification is not significant, but for $s = 1$ it completely reconciles the theoretical and experimental values of C .

It is not necessary to the theory that the elementary magnets should be as individual as has been implied here. The consideration of an average value for the whole lattice of the effective number, η , of nearest neighbours is a concession to their quantum-mechanical indistinguishability (as well as being a necessary mathematical simplification). The difference between the described ordered and disordered states may be regarded as due respectively to unequal and equal probabilities for the occupation of lattice sites. However, it is necessary to the fluctuating field model that the $3d$ electrons (or vacancies) responsible for ferromagnetism are too tightly bound to be considered as an electron band shared by the lattice as a whole.

Fluctuations of intermolecular field are inherent in the model described. They have not been given specific consideration because they become important only in the region of the Curie temperature. In a ferromagnetic in which attractive ordering occurs at low temperatures in the manner described in Section IV, the temporary formation of small coplanar groups of elementary magnets will occur even above the Curie point. This can be described either as fluctuations or as the preservation of some local order. However, when spontaneous magnetization is small, the ordering by ionic repulsion, rejected at low temperatures, may still be significant.

VI. REFERENCES

- ARROTT, A. (1955).—Carnegie Inst. Tech. Dept. of Physics, Third Quarterly Report, Ferromagnetism Research.
- BLOCH, F. (1932).—*Z. Phys.* **61**: 206.
- DYSON, F. J. (1956a).—*Phys. Rev.* **102**: 1217.
- DYSON, F. J. (1956b).—*Phys. Rev.* **102**: 1230.
- NÉEL, L. (1932).—*Ann. Phys., Paris* **18**: 5.
- NÉEL, L. (1934).—*J. Phys. Radium* **5**: 104.
- NÉEL, L. (1940).—"Le Magnetism II." (Strasbourg Conference.) p. 65. (Inst. intern. cooperation intellectuelle: Paris.)
- RUSHBROOKE, G. S., and WOOD, P. J. (1955).—*Proc. Phys. Soc. Lond. A* **68**: 1161.
- STACEY, F. D. (1955).—*Canad. J. Phys.* **33**: 661.

ON THE SEISMOLOGICAL ASPECTS OF THE 1954 HYDROGEN BOMB EXPLOSIONS

By T. N. BURKE-GAFFNEY* and K. E. BULLEN†

[Manuscript received April 29, 1958]

Summary

Tentative conclusions previously drawn from an analysis of seismic readings of four 1954 hydrogen bomb explosions are re-examined in the light of source data subsequently released on these explosions. The released data show that the authors' earlier computed origin-times for the four explosions were correct within 0.0, 0.4, 0.7, and 0.1 sec respectively. The re-examination shows that the J.B. *P* tables need a correction of -2.2 ± 1.0 sec for surface epicentres in the mid Pacific and recordings at continental stations. It is confirmed that any difference between the *P* travel-times from Bikini to Australia and Bikini to the United States is not much more than $\frac{1}{2}$ sec. The authors' previous inferences on the velocities of air waves from the explosions remain undisturbed. The re-examination confirms the occurrence of diffracted *PKP* waves in front of the 142° caustic, and confirms that these diffracted waves arrive at times significantly earlier than *PKIKP* waves.

I. INTRODUCTION

Three papers (Burke-Gaffney and Bullen 1957; Bullen and Burke-Gaffney 1957, 1958), previously published on the 1954 hydrogen bomb explosions, will be referred to as papers I, II, and III. In these papers, various inferences have been drawn from seismic records of the explosions, but the inferences were all made without knowledge of data at the explosion sources. The United States Atomic Energy Commission has now released the source data, so that the reliability of the inferences can be checked. The purpose of the present paper is to carry out this check.

The officially released (G.M.T.) origin-times of the four explosions are :

- (i) 1954 Feb. 28^d 18^h 45^m 0.0^s
- (ii) 1954 Mar. 26^d 18^h 30^m 0.4^s
- (iii) 1954 Apr. 25^d 18^h 10^m 0.7^s
- (iv) 1954 May 4^d 18^h 10^m 0.1^s

These figures show our previously estimated origin-times to have been correct within an average of 0.3 sec. This is in fact the order of accuracy that we had surmised during our calculations, and the corrections needed to the origin-times are, by themselves, too small to affect any of our previous inferences.

Our original calculations, however, rested on the postulate that the centre of each of the four explosions was the same as that for the Bikini explosion of

* Riverview College Observatory, Riverview, N.S.W.

† Department of Applied Mathematics, University of Sydney.

1946 July 24, namely, $11^{\circ} 35' \text{ N.}$, $165^{\circ} 30' \text{ E.}$ The officially released coordinates differ from these sufficiently to warrant our carrying out a re-calculation. The official coordinates are:

- (i) $11^{\circ} 41' 27'' \text{ N.}$, $165^{\circ} 16' 25'' \text{ E.}$
- (ii) $11^{\circ} 41' 27'' \text{ N.}$, $165^{\circ} 16' 23'' \text{ E.}$
- (iii) $11^{\circ} 39' 59'' \text{ N.}$, $165^{\circ} 23' 14'' \text{ E.}$
- (iv) $11^{\circ} 39' 56'' \text{ N.}$, $165^{\circ} 23' 13'' \text{ E.}$

As will be seen, our principal inferences are in fact well sustained in the re-calculations, but a few minor numerical changes are entailed.

Taking the above latitudes to be geographic, the corresponding geocentric direction-cosines are found to be:

- (i) and (ii): -0.94732 , 0.24900 , 0.20136
- (iii) and (iv): -0.94789 , 0.24714 , 0.20092 .

From these direction-cosines, we recalculated the distances of the recording stations, and then using the correct origin-times obtained revised residuals in place of those given in Table 2 of paper I. The Jeffreys-Bullen tables were again used, and ellipticity corrections were applied.

II. READINGS AT DISTANCES LESS THAN 120°

Table 1 below gives the results for stations at distances up to 120° . In forming this table, the arrival-time data set down in Table 1 of paper I were used with the sole exception of Baguio, for which revised readings had been set down in paper III. In Table 1, Δ_1 denotes the distance from the centres of shocks (i) and (ii), and Δ_2 from shocks (iii) and (iv).

The standard four-figure direction-cosines of observatories were used in computing Δ_1 and Δ_2 , so that the calculated values may be in error up to $0^{\circ}.03$. The recorded arrival-times at stations were in general available only to the nearest second, but computed travel-times, ellipticity corrections, etc. were estimated to 0.1 sec in order to avoid introducing additional errors. Thus fluctuations at least up to 0.7 sec are to be expected between individual residuals in Table 1. The entries under "mean residual" relate only to P or PKP readings.

Comparison with the residuals obtained in paper I shows that the new residuals are even slightly more self-consistent than the old. The residuals remain predominantly negative and are actually numerically a little larger than those of paper I. All the broad conclusions of paper I are confirmed. The detailed conclusions are these:

(a) The mean residual for P readings at all stations up to Kiruna is -2.2 ± 1.0 sec. This is compatible with the result -1.8 ± 0.8 sec obtained by Bullen (1948) from readings of the 1946 Bikini explosion, and confirms the fact that, for a surface epicentre in the mid Pacific, the J.-B. P tables need a negative correction of more than a second for travel-times to continental stations.

(b) The mean residuals at Riverview, Brisbane, Pasadena, and Fayetteville have all been increased from paper I. But the mean difference between the

Australian and the United States stations remains only 0.5 sec. Thus it is confirmed that the differences in travel-times from Bikini to Australia and from Bikini to the United States differ, if at all, by not much more than a half-second.

(c) The four readings at Stuttgart are very consistent and indicate a correction of -1.0 ± 0.4 sec to the J.-B. PKP table for a mid Pacific surface focus. The two readings at Tananarive do not support a need for this correction, but it is likely that the first onsets at Tananarive were too small to be recorded.

(d) The conclusions in paper I on the velocities of air waves from the explosions remain unaffected.

TABLE I
P AND PKP RESIDUALS AT DISTANCES UP TO 120°

Station	Δ_1	Δ_2	(i) (sec)	(ii) (sec)	(iii) (sec)	(iv) (sec)	Mean Residual (sec)
Noumea ..	$33^\circ.72$	$33^\circ.69$	-0.4			-0.2	-0.3
Matsushiro	$34^\circ.73$	$34^\circ.82$	+4.3(S)	-1.4			-1.4
Brisbane ..	$40^\circ.71$	$40^\circ.72$	-1.1	-1.5	-1.9	-1.3	-1.5
Baguio ..	$43^\circ.54$	$43^\circ.66$	-1.6	-2.0	-3.3	-2.7	-2.4
Riverview	$47^\circ.22$	$47^\circ.23$	-1.7	-1.1	-1.5	-1.9	-1.5
Kaimata ..	$54^\circ.15$	$54^\circ.12$				-1.3	-1.3
Victoria ..	$69^\circ.01$	$68^\circ.95$	-4.5	-4.9		-2.3	
Pasadena	$72^\circ.46$	$72^\circ.38$	-1.6	-2.0	-2.0	-2.4	-2.0
Fayetteville	$91^\circ.54$	$91^\circ.47$	-2.3	-1.7			-2.0
Quetta ..	$91^\circ.26$	$91^\circ.37$	-3.1		-2.3	-2.7	-2.7
Kiruna ..	$96^\circ.75$	$96^\circ.80$				-3.6	-3.6
Uppsala ..	$104^\circ.11$	$104^\circ.17$				+5.1	+5.1
Ksara ..	$113^\circ.94$	$114^\circ.06$		0 (PP)*			
Stuttgart	$116^\circ.19$	$116^\circ.25$	-1.5	-0.9	-1.3	-0.5	-1.0
Tananarive	$119^\circ.75$	$119^\circ.85$		+0.2		+0.4	+0.3

* If the reading at Ksara were taken as $48^m 42^s$ instead of $49^m 42^s$, the reading would agree with PKP with a residual of -0.1 sec.

III. READINGS IN THE RANGE $137^\circ < \Delta < 142^\circ$

In papers II and III, attention was drawn to what, on the calculations of paper I, appeared to be abnormally early readings at all three stations in the range $137^\circ < \Delta < 142^\circ$. It was suggested that the early readings related to diffracted PKP waves associated with the caustic at 142° . The newly released source data referred to earlier in this paper enable us to confirm this conclusion.

In Table 2 below, we give the revised results for the three stations. The distances Δ_1 and Δ_2 are given in brackets below the names of stations. For Pretoria and Kimberley, the revised residuals correspond to the readings, both first onsets and certain later phases, made by Dr. Hales and referred to in paper III. For Tamanrasset, the revised residuals correspond to the arrival-times given in paper I.

In Table 2, the means of the first arrivals at the three stations are -11.0, -8.3, -6.1 sec respectively, as against -11.1, -8.2, -6.1 sec, obtained in

paper I. Means of the late arrivals at Pretoria and Kimberley are likewise much the same as before.

Thus the conclusions of paper III on the existence of diffracted *PKP* waves in front of the 142° caustic, arriving at times significantly earlier than *PKIKP* waves, are totally substantiated. The official release of the source data of the explosions has thus confirmed the importance of the explosions in adding a final link in the chain of evidence on the existence of the Earth's inner core.

TABLE 2
RESIDUALS AT STATIONS BETWEEN DISTANCES OF 137° AND 142°

Station	(i) (sec)	(ii) (sec)	(iii) (sec)	(iv) (sec)
Pretoria ($137^\circ 22$, $137^\circ 31$)	-11.4, -2.9, -0.1	-11.0, +1.2		-10.6, +1.4
Kimberley ($139^\circ 45$, $139^\circ 54$)	-7.8, -1.4, +1.3	-8.2, +0.8	-6.7	-10.7, +1.8
Tamanrasset ($140^\circ 40$, $140^\circ 48$)	-6.2	-8.6	-5.1	-4.6

The authors would like to record the service which the United States Atomic Energy Commission has rendered to seismology in releasing the source data on the explosions here discussed. The release has enabled the previous tentative conclusions to become well-established conclusions. And the importance of controlled explosions as a means of studying the Earth's deep interior is confirmed.

IV. REFERENCES

- BULLEN, K. E. (1948).—*Nature* **161**: 62.
 BULLEN, K. E., and BURKE-GAFFNEY, T. N. (1957).—*Nature* **180**: 49.
 BULLEN, K. E., and BURKE-GAFFNEY, T. N. (1958).—*Geophys. J. R. Astr. Soc.* **1**: 9.
 BURKE-GAFFNEY, T. N., and BULLEN, K. E. (1957).—*Aust. J. Phys.* **10**: 130.

A STUDY OF "SPREAD- F " IONOSPHERIC ECHOES AT NIGHT AT BRISBANE

IV. RANGE SPREADING

By H. C. WEBSTER*

[Manuscript received May 5, 1958]

Summary

In the course of investigations of satellite echoes from the F region of the ionosphere, it was noted that the F and E_s traces recorded at night, on $h'f$ equipment at frequencies well below vertical, are broader than anticipated and tend to change in a characteristic manner as the gain of the receiver is lowered. In this paper, a quantitative explanation of these phenomena is elaborated, based on the postulate of a "rough" ionosphere.

This theory leads to a method whereby, from the swept-gain $h'f$ records, estimates of roughness index can be formed. These estimates compare satisfactorily, on a statistical basis, with estimates by other methods. The theory is extended to the case of multiple-hop reflections, and to the satellite traces; general agreement with experiment is found. Evidence is presented that the ionosphere appears rougher when transmitter and receiver are adjacent than when they are widely separated, and a tentative explanation is suggested. From the roughness indices, the relative intensities of Z - and O -mode F echoes for Brisbane are computed and the rare appearance of Z -traces on Brisbane records is satisfactorily explained.

I. INTRODUCTION

The fixed-frequency $h'f$ recorder described by McNicol, Webster, and Bowman (1956) in Part I of this series of papers gives F traces which in almost all cases are substantially broader than would be anticipated from the duration of the pulse, even when allowance is made for the finite bandwidth of the receivers used and for the finite size of spot of the cathode-ray tube. In the swept-gain records discussed by those authors it is found that, as the gain decreases, the trace contracts until, just before extinction, it is no wider than would be anticipated on the basis of pulse duration, receiver response, and C.R.T. spot size (cf. Plate 1, Fig. 1). Occasionally, at a somewhat higher gain, the trace divides, indicating the presence of a "satellite" (McNicol, Webster, and Bowman, loc. cit.) which was not separated from the main trace when the gain was maximal (Plate 1, Fig. 2). If we restrict consideration to those swept-gain records in which there is no such unresolved satellite, it is found that the patches on the swept-gain records tend to take up a shape approximating to a trapezium (cf. Fig. 1). Individual patches depart from this shape, but the average shape of a number of successive patches resembles it reasonably well. The angle BAD (Fig. 1) can change markedly from one hour to the next, but the angle ABC is more nearly constant. The distance BE (where angle CDE is made equal to BCD) usually

* Physics Department, University of Queensland, Brisbane.

SPREAD-F IONOSPHERIC ECHOES AT NIGHT AT BRISBANE. IV

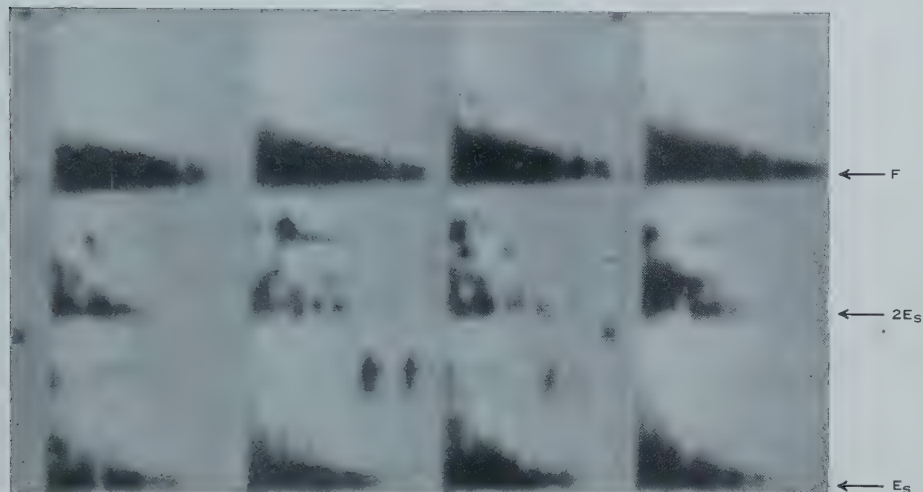


Fig. 1.—Example of swept-gain patches (March 24, 1955). Upper trace, F -region echo. Lower traces, E_s and double-hop E_s . Duration of swept-gain cycle 2 min.

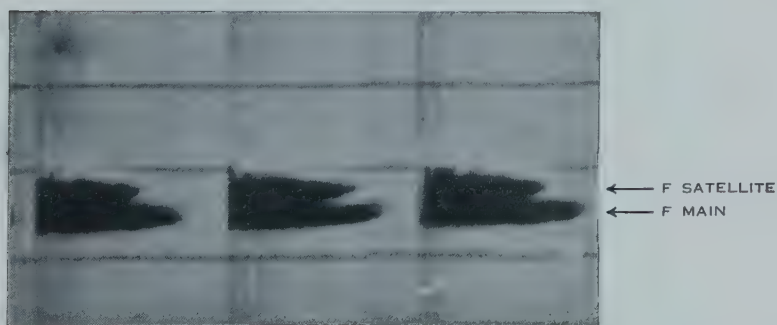
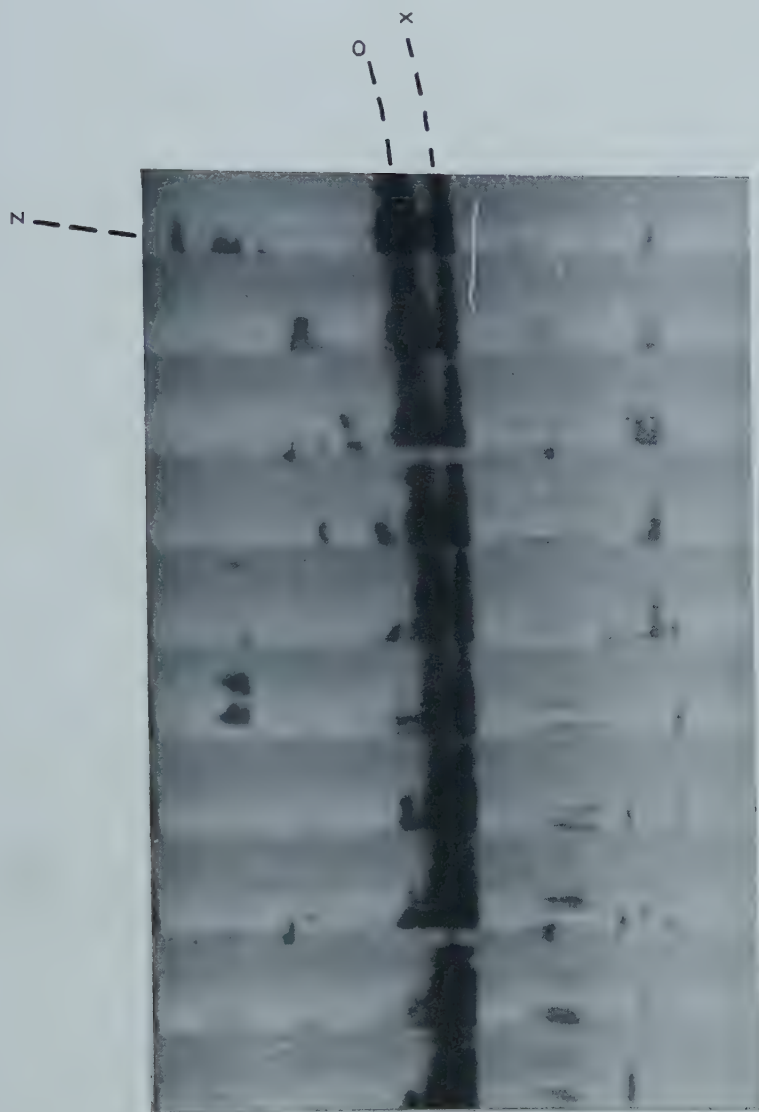


Fig. 2.—Example of swept-gain patches (December 3, 1955), showing F satellite resolved at low gain, unresolved at high (E_s not recorded on this occasion).

SPREAD-F IONOSPHERIC ECHOES AT NIGHT AT BRISBANE. IV

Example of Z-ray recorded on swept-gain $h'f$ record in Brisbane, 2.28 Mc/s.

corresponds, roughly, to the trace width calculated from pulse duration and receiver characteristics. Examination of swept-gain records of double-hop and triple-hop echoes shows similar effects and such effects also appear frequently (a) on the records of "satellites" and (b) on records obtained when there is 100 km separation between transmitter and receiver.

In the present paper a tentative theory of this broadening is worked out, on the basis of a uniformly rough ionosphere, employing assumptions similar to those used by Briggs and Phillips (1950) in their study of the fading of reflected pulses, and quantitative experimental data obtained from swept-gain records are then discussed in relation to this theory. While this paper was in preparation, Moore and Williams (1957) have published an analysis of a somewhat analogous problem (ground scatter of radio emissions from an airborne transmitter). Their results appear to be generally consistent with those given here, but a detailed comparison is scarcely feasible.

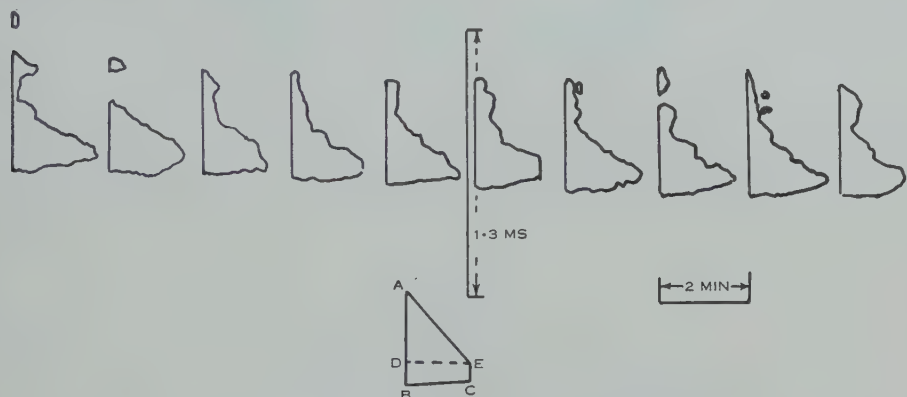


Fig. 1.—Outlines of swept-gain patches and their idealized trapezoidal form.

II. GENERAL THEORETICAL CONSIDERATIONS

Briggs and Phillips (1950) show that the energy which is returned to a transmitting station by a small element of area of the ionosphere subtending a small solid angle $d\Omega$ at the transmitter and located at an angular distance θ from the zenith is given by

$$dW = \kappa \cos^n \theta d\Omega,$$

where κ is a constant depending on the transmitter power, antenna gain, etc. and n is an exponent dependent on the state of roughness of the ionosphere, antenna patterns, etc. This conclusion is based on reasonable assumptions and leads to a consistent explanation of fading data.

They also assume that the total energy received from the whole ionosphere is obtained by adding the power contributions from the different elements. This amounts to assuming that the signals scattered from different areas have random differences of phase. In the case of specular reflection, this cannot be true, but in that case the effective value of n is so large that no discrepancy arises. The smaller the value of n the more nearly the ionosphere approaches

complete roughness. Values of n below 3 are impossible and values below 10 very improbable, under the conditions of observation used in the Brisbane experiments.

It will be assumed that the dispersion of the ionosphere is so small that it does not affect the pulse width. This is not valid near the penetration frequency.

On the basis of these assumptions and assuming also an ionosphere devoid of major irregularities or curvatures, but having such minor irregularities as

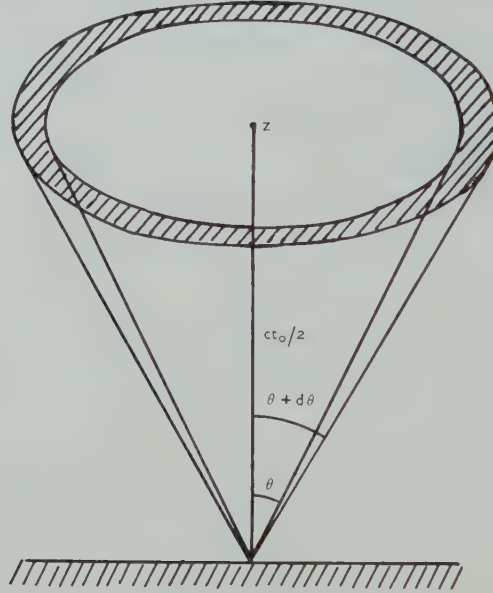


Fig. 2.—Scattering from a rough ionosphere, transmitter and receiver contiguous. Z is point on reflecting surface vertically above station.

to lead to a finite value for n , it follows immediately that the energy received back in a time dt from a strip of the ionosphere bounded by cones of semi-angles θ and $\theta + d\theta$ (Fig. 2) must be proportional to $F(\theta)$ where

$$F(\theta)d\theta = 2\pi \cos^n \theta \sin \theta d\theta. \quad \dots\dots\dots (1)$$

The time taken for this radiation to travel from the station and back is given by

$$t = t_0 \sec \theta,$$

where t_0 is the time taken for vertically reflected radiation.

If the ionosphere behaved as a specular reflector, the shape of the received pulse would depend only on

- (a) the shape of the pulse emitted by the transmitter ;
- (b) distortion (due to finite bandwidth) of the receiver.

It is convenient to lump these two effects together, assuming a distortionless receiver with a transmitter pulse adjusted in shape to correspond to the pulse as it

appears in the receiver when specular reflection has occurred. This procedure would appear to be valid if the receiver is linear. We shall represent the shape by the arbitrary power function $P(\tau)$, i.e. $P(\tau)$ is the radiated power at time τ after the beginning of the pulse.

At a time t after the start of the pulse, where $t > t_0$, radiation which originated at the start of the pulse will be arriving from a ring on the ionosphere at an angle θ_1 , given by

$$t = t_0 \sec \theta_1.$$

At the same time t , radiation will also be arriving after reflection from more nearly zenithal ionospheric rings; this radiation will have been emitted at a later time τ within the pulse. In fact, all parts of the pulse, up to a time τ_1 , will be contributing, where

$$\tau_1 = t - t_0,$$

and corresponds therefore to zenithal reflection. Still later parts of the pulse do not contribute to the radiation being received at time t . If $t - t_0 > \tau_0$, the whole pulse contributes.

The energy ΔW returning in an interval Δt at time t from an infinitesimal ring enclosed between cones of angles θ and $\theta + d\theta$ ($\theta < \theta_1$) is therefore proportional to the energy emitted during the corresponding interval τ to $\tau + \Delta\tau$, where

$$\begin{aligned}\tau &= t - t_0 \sec \theta, \\ \tau + \Delta\tau &= t + \Delta t - t_0 \sec \theta.\end{aligned}$$

Since it is also proportional to $F(\theta)d\theta$ we may write

$$\Delta W = kP(\tau)\Delta\tau F(\theta)d\theta = kP(t - t_0 \sec \theta)\Delta t F(\theta)d\theta,$$

where k is a constant.

The total rate of return of energy at time t may be obtained by adding the contribution of the several rings and thus becomes

$$I = k \int_0^{\theta_1} P(t - t_0 \sec \theta) F(\theta) d\theta. \quad \dots\dots\dots (2)$$

The shape of the swept-gain patches is determined by the dependence of $\log I$ on t . In order to evaluate the integral we must make some assumptions regarding the form of $P(\tau)$.

III. IDEAL PULSE

It is instructive to consider first the idealized case where the pulse is strictly rectangular in shape. We shall take its duration as τ_0 . Thus

$$\begin{aligned}P(t - t_0 \sec \theta) &= P & (t_0 + \tau_0 > t > t_0), \\ &= 0 & (t > t_0 + \tau_0).\end{aligned}$$

Two cases arise, according to whether $t - t_0 \geq \tau_0$, i.e. $\tau_1 \geq \tau_0$.

Case A. $0 < \tau_1 < \tau_0$. Then the function $P(t - t_0 \sec \theta)$ is the same for all values of θ less than θ_1 , and

$$I = 2\pi k P \int_0^{\theta_1} \cos^n \theta \sin \theta \, d\theta \\ = \frac{2\pi k P}{n+1} \left\{ 1 - \left(\frac{t_0}{t} \right)^{n+1} \right\}. \quad \dots\dots\dots (3)$$

The quantity within the braces increases from zero (for $t = t_0$) to a value which, if n is large, can be near unity (for $t = t_0 + \tau_0$).

Case B. $\tau_1 > \tau_0$. In this case vertically reflected radiation would have had to originate after the end of the pulse. In fact therefore no radiation is being received from within a cone of semi-angle θ_2 , where

$$t_0 \sec \theta_2 = t - \tau_0.$$

θ_2 is the angle corresponding to the end τ_0 of the emission. Thus in this case

$$I = 2\pi k P \int_{\theta_2}^{\theta_1} \cos^n \theta \sin \theta \, d\theta \\ = \frac{2\pi k P}{n+1} \left\{ \left(\frac{t_0}{t - \tau_0} \right)^{n+1} - \left(\frac{t_0}{t} \right)^{n+1} \right\}. \quad \dots\dots\dots (4)$$

The function has its maximum value when

$$t = t_0 + \tau_0,$$

this value being

$$I_m = \frac{2\pi k P}{n+1} \left\{ 1 - \left(\frac{t_0}{t_0 + \tau_0} \right)^{n+1} \right\}.$$

Thus

$$\frac{I}{I_m} = \frac{\{t_0/(t - \tau_0)\}^{n+1} - (t_0/t)^{n+1}}{1 - \{t_0/(t_0 + \tau_0)\}^{n+1}}. \quad \dots\dots\dots (5)$$

The shapes of the swept-gain patches indicate that n is usually a large number. Since in the records discussed τ_0/t_0 is of the order 0.05, it is clear that the first terms in the numerator and denominator far exceed the second and the latter can thus be neglected. Thus, to a good approximation,

$$\frac{I}{I_m} = \{t_0/(t - \tau_0)\}^{n+1}, \quad \dots\dots\dots (6) \\ \log I = \log I_m + (n+1) \log t_0 - (n+1) \log (t - \tau_0).$$

The slope of the upper edge of the swept-gain patch (i.e. beyond its maximum) should thus correspond to S_1 , where

$$S_1 = \frac{1}{\partial (10 \log I) / \partial t} = \frac{t - \tau_0}{10 \log e} \cdot \frac{1}{n+1} \quad \dots\dots\dots (7)$$

S_1 would thus vary sufficiently slowly with t to be sensibly constant over the width of a normal trace. Note that it is proportional to $1/(n+1)$ and thus the measurement of the slope allows n to be determined.

IV. TRIANGULAR PULSE

The actual shape of the pulse, received under specular reflection conditions, is far from rectangular. In appearance it approximates roughly to an error function, but as this function does not yield conveniently integrable expressions, calculations have been made assuming a triangular pulse, defined by

$$\begin{aligned} P(\tau) &= b\tau/\tau_0 & (0 < \tau < \tau_0/2), \\ P(\tau) &= b(\tau_0 - \tau)/\tau_0 & (\tau_0/2 < \tau < \tau_0), \\ P(\tau) &= 0 & (\tau > \tau_0). \end{aligned}$$

In this case three cases arise, corresponding to these three conditions, namely:

Case (a). $0 < \tau_1 < \tau_0/2$.

$$I = \frac{2\pi kb}{n+1} \left[\frac{t-t_0}{\tau_0} - \frac{1}{n} \cdot \frac{t_0}{\tau_0} \left\{ 1 - \left(\frac{t_0}{t} \right)^n \right\} \right]. \quad \dots\dots\dots (8)$$

Case (b). $\tau_0/2 < \tau_1 < \tau_0$.

$$I = \frac{2\pi kb}{n+1} \left[\frac{t_0 - t + \tau_0}{\tau_0} + \frac{1}{n} \frac{t_0}{\tau_0} \left\{ 1 - 2 \left(\frac{t_0}{t - \tau_0/2} \right)^n + \left(\frac{t_0}{t} \right)^n \right\} \right]. \quad \dots\dots (9)$$

Case (c). $\tau_1 > \tau_0$.

$$I = \frac{2\pi kb}{n(n+1)} \frac{t_0}{\tau_0} \left(\frac{t_0}{t - \tau_0} \right)^n \left\{ 1 - 2 \left(\frac{t - \tau_0}{t - \tau_0/2} \right)^n + \left(\frac{t - \tau_0}{t} \right)^n \right\}. \quad \dots (10)$$

To examine the general character of these results we shall again have recourse to approximation. If n is a large quantity (specifically if $n\tau_0 \gg t$) then the results reduce, respectively, to

$$(a) \quad \frac{2\pi kb}{n+1} \left[\frac{t-t_0}{\tau_0} - \frac{1}{n} \frac{t_0}{\tau_0} \right]. \quad \dots\dots\dots (11)$$

$$(b) \quad \frac{2\pi kb}{n+1} \left[\frac{t_0 - t + \tau_0}{\tau_0} + \frac{1}{n} \frac{t_0}{\tau_0} \right]. \quad \dots\dots\dots (12)$$

(This approximation is invalid if t is very near $t_0 + \tau_0/2$.)

$$(c) \quad \frac{2\pi kb}{n+1} \frac{1}{n} \frac{t_0}{\tau_0} \left(\frac{t_0}{t - \tau_0} \right)^n. \quad \dots\dots\dots (13)$$

It will be noted that for the upper edge of the swept-gain patch we now obtain

$$S_1 = \frac{1}{\partial (10 \log I) / \partial t} = - \frac{t - \tau_0}{10 \log e} \cdot \frac{1}{n}, \quad \dots\dots\dots (14)$$

which is not distinguishable (for large n) from the result obtained previously. This suggests that (14) may be approximately valid also for the error function form of $P(\tau)$.

As an example of the effect on the shape of swept-gain patches of the (effective) emitted pulse shape, the dependence of $\log I$ on $t-t_0$ has been computed for the following conditions :

$$t_0 = 10^{-3} \text{ sec}, \quad n = 500,$$

(a) Rectangular $\tau_0 = 5 \times 10^{-5} \text{ sec},$
 (b) Triangular $\tau_0 = 10^{-4} \text{ sec},$

and the results depicted in Figure 3 (with time-zero displaced appropriately).

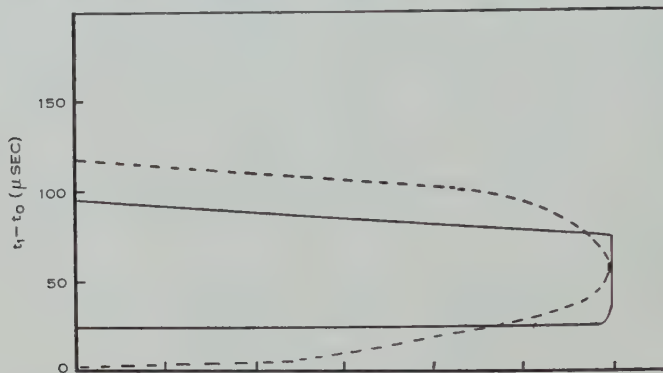


Fig. 3.—Theoretical shape of swept-gain patches, $n=500$, $t_0=10^{-3} \text{ sec}$, $\tau_0=5 \times 10^{-5} \text{ sec}$ (rectangular pulse), 10^{-4} sec (triangular pulse).

V. DOUBLE-HOP REFLECTIONS

A complete treatment, along the same lines, of double-hop reflections is scarcely feasible because of the wide variety of possible paths. Scattering can occur, not only at each reflection from the ionosphere, but also at the ground (e.g. Dieminger 1951). It seemed, however, worth while to consider briefly

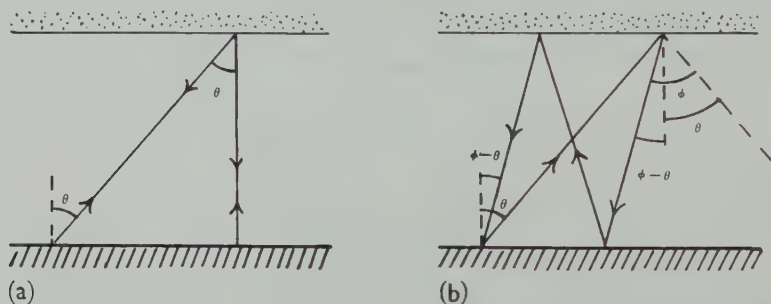


Fig. 4.—Scattering from a rough ionosphere in double-hop reflections (specular reflection at ground). (a) Outward and inward paths coincident, two scattering processes; (b) outward and inward paths different, one scattering process.

two limiting cases, in both of which specular reflection from the ground is assumed. In the case represented in Figure 4 (a), radiation is scattered (non-specularly) twice from the ionosphere, in the case represented in Figure 4 (b) once only.

Comparing the first of these cases with the single-reflection case, the radiation is scattered twice through an angle θ instead of once through 2θ . Thus the scattering factor (cf. Briggs and Phillips, loc. cit.) becomes $(\cos \theta \cos^q \theta)^2$ instead of $\cos \theta \cos^q 2\theta$ and the overall intensity is proportional to

$$\cos^{a+2q+8}\theta \text{ instead of } \cos^{a+4q+6}\theta$$

where $a=t+r$, t, r representing the properties of transmitter and receiver antenna systems in Briggs and Phillips' notation.

In the experiments to be described, a is small (of the order of 2 or 3), while the shape of the patches indicates a large value of n . Thus q is the preponderant term in the index.

To forecast the overall slope of the swept-gain patch we have to take into account also the fact that in this case

$$t=t_0(1+\sec \theta_1)$$

(with t_0 corresponding to a single vertical reflection).

Applying the same arguments as in Section III, and also introducing the approximation

$$t-\tau_0=2t_0,$$

it is possible to show that the slope of the swept-gain patch, in the double-hop case, should correspond to

$$S_2 = -\frac{t_0}{10 \log e} \cdot \frac{1}{n_2}, \quad \dots \dots \dots (15)$$

where $n_2=a+2q+9$.

For the single-hop case, using the corresponding approximation ($t-\tau_0=t_0$) we obtain

$$S_1 = -\frac{t_0}{10 \log e} \cdot \frac{1}{n_1},$$

where $n_1=a+4q+7$. Thus

$$S_2/S_1=n_1/n_2=2 \text{ (approx.)}, \quad \dots \dots \dots (16)$$

if q is large.

In the case represented in Figure 4 (b) the important factor in the expression for the intensity of the scattered wave is $\cos^q \varphi$, while we have

$$t=\frac{t_0}{2} \sec \theta + \frac{3t_0}{2} \sec (\varphi - \theta).$$

A numerical calculation indicates that, if we set

$$\sec \chi = t/2t_0,$$

then, roughly,

$$\begin{aligned} \varphi &= 2.25\chi, \text{ and thus, if } q \text{ is large,} \\ \cos^q \varphi &= \cos^{5q}\chi. \end{aligned}$$

In view of the smallness of a , it is permissible to use χ in place of θ and $(\varphi - \theta)$ in the remaining factors and write

$$dW \propto \cos^{a+5q+6} \chi.$$

For the slopes of the swept-gain patches we now have

$$S'_2 = \frac{-t_0}{10 \log e} \frac{1}{n'_2}, \dots\dots\dots (17)$$

where $n'_2 = a + 5q + 6$,

and thus

$$S'_2/S_1 = n_1/n'_2 = 0.8 \text{ (approx.)}. \dots\dots\dots (18)$$

Thus the ratio of slopes is expected to lie between 0.8 and 2.0 dependent on the relative importance of the mechanisms responsible.

VI. THE BROADENING OF RESOLVED SATELLITE TRACES

According to the views elaborated by McNicol and Webster (1956), the presence of a satellite trace indicates the existence of a large-scale irregularity, often of considerable extent, in the F layer. Although this irregularity can be regarded, as a first approximation, as a straight wave-front, there is evidence that minor irregularities are usually present along its length and also that the ionospheric surfaces on either side of it are not specular reflectors.

Since the conditions specified in Section II are no longer valid, we might expect a departure from trapezoidal shape (for the main trace) when satellites are present. Moreover, in the case of the satellites themselves, since the effective scattering surface is now a narrow strip rather than the complete ionospheric plane (strictly, sphere) we might expect the contributions of more remote parts to be less important, i.e. in the swept-gain patch we might expect a smaller slope of the upper edge than for the main trace.

VII. COMPARISON OF THEORY WITH EXPERIMENT

In principle, it should be possible to subject the theory to direct check, by deducing from the slope of the swept-gain patches the anticipated correlation coefficient between fading records on spaced receivers and comparing these values with those obtained in fading experiments. An attempt was made to do this using the fading records obtained by Burke and Jenkinson (1957) in measurements of ionospheric drifts. The results were indefinite; in some cases the calculated correlation coefficient was much higher than the experimental, in other cases much lower. This discrepancy may have been due partly to the effect of the automatic gain control in the fading receivers and partly to inadequacy of the fading samples obtained.

Some confirmation of the accuracy of the conclusions is, however, available from a comparison of the spread of values of n for F reflections determined from the swept-gain patches in a random sample of records, with the spread of values found by Briggs and Phillips (*loc. cit.*) on almost the same frequency. The results of the present investigation are shown in histogram form in Figure 5

(which should be compared with Figure 8 of Briggs and Phillips' paper). For convenience in comparison, Figure 5 represents, not n , but θ_0 , defined by

$$(\cos \theta_0)^n = \frac{1}{4},$$

from which, if θ_0 is small, $\theta_0 = 95 \cdot n^{-\frac{1}{2}}$ degrees.

Figure 5 shows a modal value slightly greater than obtained by Briggs and Phillips. This discrepancy could well be due to the presence of unresolved satellites leading to over-estimates of the slope of the swept-gain patches. Figure 5 includes data obtained from records showing resolved satellites, as well as those showing no resolved satellites. No systematic relationship appears to exist between θ_0 -values and the presence or absence of resolved satellites.

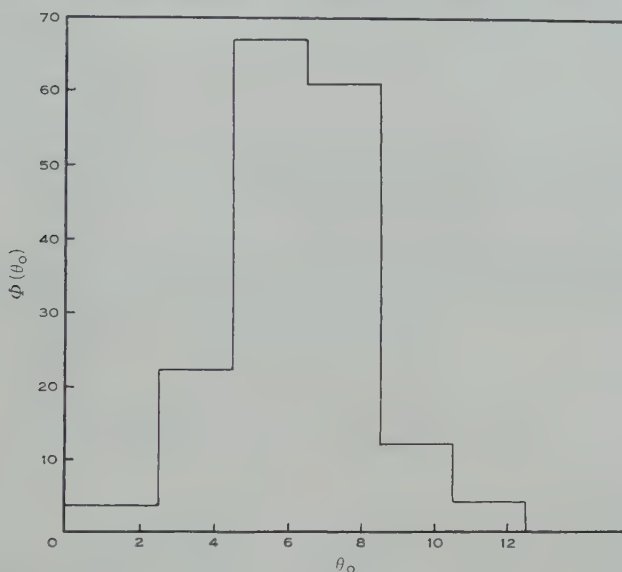


Fig. 5.—Histogram showing distribution of θ_0 values among 160 cases. $\Phi(\theta_0)$ represents number of cases having θ_0 values between limits shown.

The E_s traces have been analysed similarly in a number of cases. As would be expected from the nature of the night-time E_s layer, the mean value of θ_0 is higher than for the F (14 v. 5.9). The distribution is roughly consistent with that obtained by Briggs and Phillips for night-time E_s at Cambridge on a similar frequency.

Two-hop F reflections are recorded on some swept-gain films. One film has been examined in detail. In Figure 6, values of θ'_0 corresponding to $2n_2$, where n_2 is obtained from the slope of $2F$ patches using equation (15), have been plotted against the corresponding values of θ_0 for the single-hop patches. The points are roughly symmetrical about the line corresponding to $n_2 = 0.6n$, and two-thirds lie between or close to the lines corresponding to $n_2 = 0.5n_1$ and $n'_2 = 1.25n_1$. The fact that the $2F$ patches conform less closely to the "ideal" shape probably accounts for some discrepancies; neglect of ground-scatter may also be a factor.

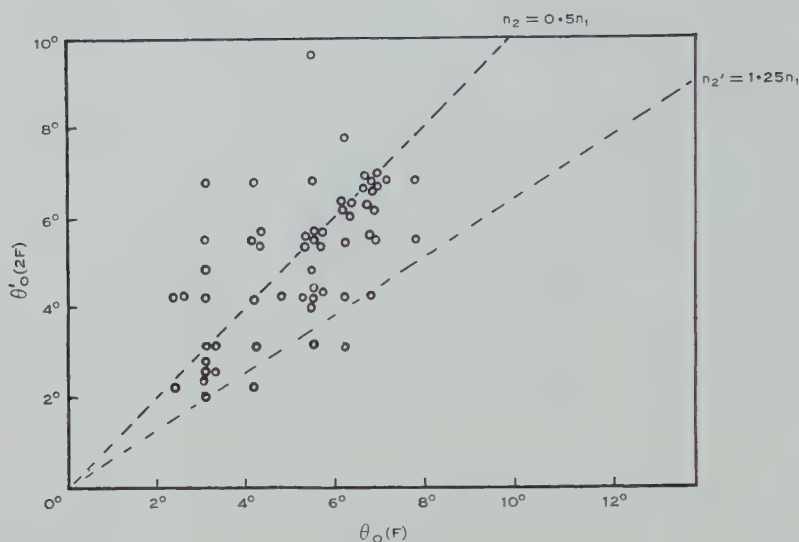


Fig. 6.—Comparison of θ'_0 deduced from double-hop F reflections (assuming $n_2=0.5n$) with θ_0 deduced from the corresponding single-hop reflection. The broken lines indicate theoretical relations for the process of Figures 4 (a) (upper line) and 4 (b) (lower line).

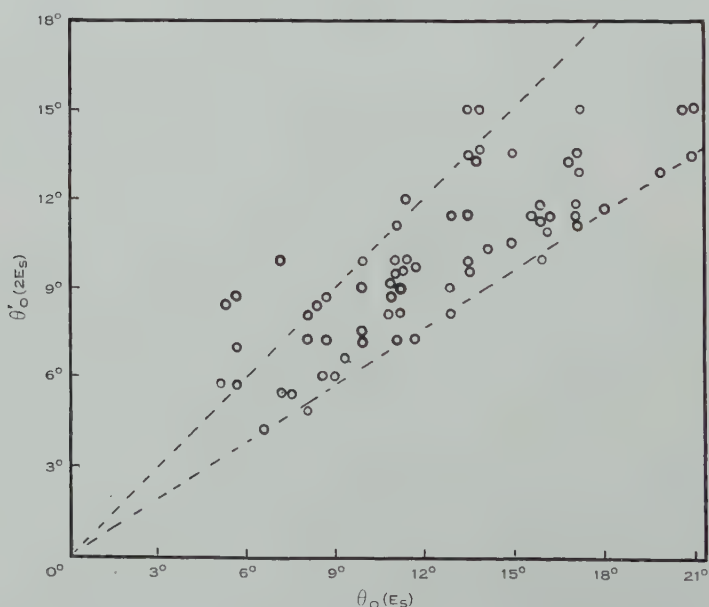


Fig. 7.—Comparison of θ'_0 deduced from double-hop E_s reflections (assuming $n_2=0.5n$), with θ_0 deduced from the corresponding single-hop reflection. The broken lines indicate theoretical relations, the upper line for the process of Figure 4 (a), the lower for that of Figure 4 (b).

Figure 7 shows, for comparison, the corresponding situation for the E_s and $2E_s$ reflections (on other films). Here rather a higher percentage of points lie between the predicted limits. A few swept-gain patches of triple-hop E_s have been examined. If we write

$$n_3 = \frac{-t_0}{\log e} \frac{\partial \log I}{\partial t}$$

and calculate the values of θ_0 corresponding to $3n_3$, the values obtained are somewhat lower than for the corresponding two-hop reflection.

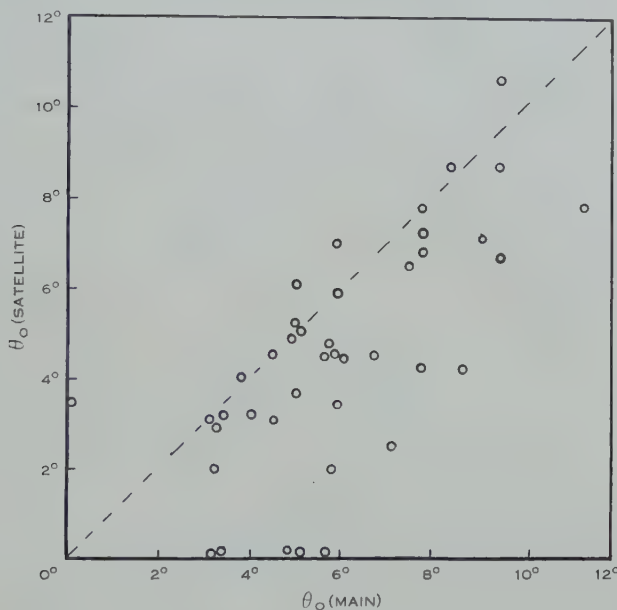


Fig. 8.—Comparison of θ_0 value for F satellite with θ_0 value for corresponding main trace.

In Figure 8 values of θ_0 determined from a “satellite” trace are plotted against θ_0 determined from the corresponding main trace. It will be noted that in the majority of cases the satellite trace gives a smaller value of θ_0 , as the theory would suggest. A few satellites without any measurable spreading ($\theta_0 \rightarrow 0$) were recorded. These were usually of rather low (< -40 dB) relative intensity and hence were probably due to particularly narrow wave fronts.

VIII. TRACE-BROADENING USING REMOTE TRANSMITTER

Some swept-gain records were made at Brisbane in which F echoes from transmissions from Toowoomba (100 km away) were recorded simultaneously with F echoes from Brisbane transmissions. In Figure 9, values of θ_R , the angle corresponding to the value of n , estimated from the swept-gain patches for Toowoomba transmissions, are plotted against corresponding values θ_A for Brisbane transmissions. In 25 per cent. of cases values were nearly equal, and

in 70 per cent. a lower value was found for the distant than for the local transmitter. Figure 10 gives some indication of the corresponding situation for E_s reflections. The adjustment of the time-base was such that the one-hop E_s reflection from the adjacent transmitter was not recorded. Values of θ' were estimated from the value of n_2 for the two-hop E_s trace, using equation (16), and compared with values of θ_R for the one-hop E_s trace from the remote transmitter. For 75 per cent. of the points, θ_R is less than θ'_A . It will be noted that if θ_A had been computed using equation (18), the difference would be still more marked.

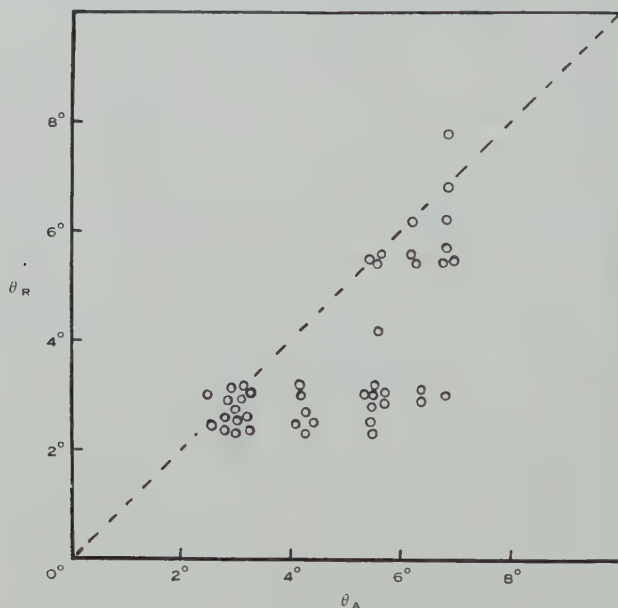


Fig. 9.—Comparison of θ_0 values from F traces arising from a remote transmitter (θ_R), with corresponding θ_0 value from a contiguous transmitter (θ_A).

These results are in line with the observations made during the experiments of McNicol, Webster, and Bowman (*loc. cit.*), and of Strohfeldt, McNicol, and Gipps (1952), that there is less evidence of (resolved) satellites when the transmitter is remote from the receiver. In searching for an explanation of the effect, it would seem necessary to exclude any mechanism involving refractive effects, since these would be strongly frequency-dependent, whereas range broadening is only slightly, if at all, dependent on frequency. Moreover, except in traces where there are, fairly clearly, imperfectly resolved satellites, the phase-path records usually show straight fringes running right across the F trace. This also, in many instances, applies to the E_s trace. A geometrical model, having approximately the required properties, consists of a flat ionosphere eroded by narrow pits (say, 10 km diameter) of approximately hemispherical shape. A hemispherical pit would reflect back along the path of incidence any ray passing

through the centre of the sphere, but rays incident on the hemispherical surface in other directions would only return to Earth after double reflection from the surface. A slight rounding-off of the edges of some of the pits would provide a small amount of "scattering" of the rays from the remote transmitter. It is worthy of note that McNicol and Webster have suggested canyons in the ionosphere in explaining some of the features of satellite traces.

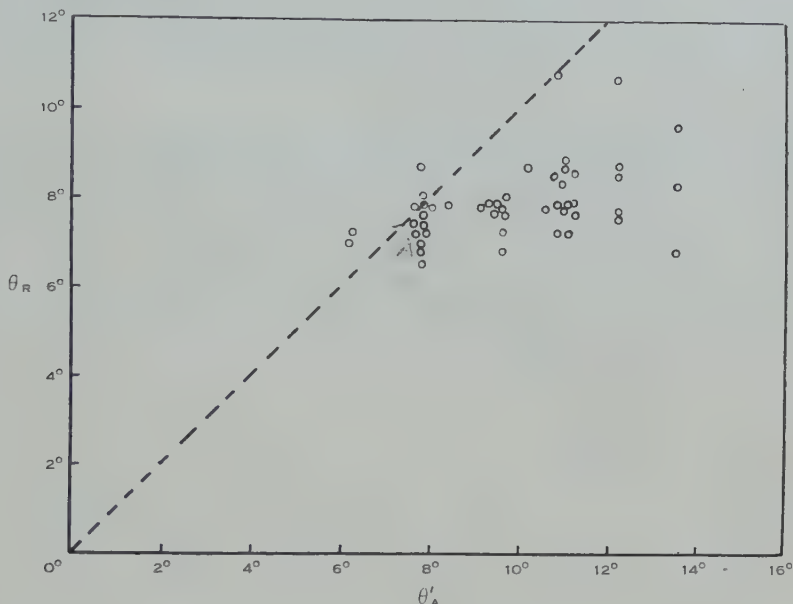


Fig. 10.—Comparison of θ_0 values deduced from E_s traces arising from a remote transmitter (θ_R) with corresponding θ'_0 values from $2E_s$ traces from a contiguous transmitter (θ'_A).

IX. APPLICATION TO Z-RAY INTENSITY

At swept-frequency stations at high latitude, a third (Z) trace frequently accompanies the O - and X -traces, the separations $f_0 - f_z$ and $f_x - f_0$ being approximately equal. This phenomenon has never been clearly recorded on swept-frequency records at Brisbane, but it has on a few occasions appeared on fixed-frequency records (cf. Plate 2) especially swept-gain and phase-path (McNicol, Webster, and Bowman, loc. cit.). Ellis (1956) has put forward a theory of the phenomenon which these records make it possible to check. He shows that for rays having directions lying within a small cone surrounding a critical direction θ_c , which at Brisbane would lie at $19^\circ 30'$, to the vertical and in the meridional plane, the O -ray would pass beyond the level at which it is usually reflected and would be reflected at a level where (in the absence of collisions) the parameters satisfy the equation

$$x = 1 + y,$$

where $x = Ne^2/m\omega^2\epsilon_v$ (using rationalized MKSA units) and $y = \omega_H/\omega$, where $\omega_H/2\pi$ is the gyro frequency.

However, this reflected radiation cannot return to the Earth's surface; it needs to be scattered back along the direction of incidence to do so. From our previous assumptions the scattered intensity in this direction would be proportional to $\cos^n \theta_c \Delta\Omega$ where $\Delta\Omega$ is the effective solid angle of the cone of rays which pass through the O reflection level. Ellis has found, experimentally, that the semi-angle to the half-power points is 0.42° . It is convenient, however, to use an equivalent cone, having full power transmitted for all directions within the cone and zero power for directions outside it. Assuming the actual distribution to be Gaussian, this gives $\Delta\theta = 0.52^\circ$ for the semi-angle of the equivalent cone, and $\Delta\Omega = \pi\Delta\theta^2 = 2.58 \times 10^{-4}$ steradian.

Since the angle $\Delta\theta$ is much smaller than the separation $\theta_1 - \theta_2$ of the limits of integration, the Z -echo should thus be of the same duration as the pulse itself and show no variation of intensity. The ratio of the intensity of the Z -ray to the maximum intensity of the O -ray should, in the absence of differences in absorption etc., be given by

$$\frac{I_z}{I_m} = \frac{\cos^n \theta_c \Delta\Omega}{\{2\pi/(n+1)\}[1 - \{t_0/(t_0 + \tau_0)\}^{n+1}]} = \frac{\cos^n \theta_c \Delta\Omega}{2\pi\tau_0/t_0} \text{ approx.,} \quad \dots\dots\dots (19)$$

$$\log \frac{I_z}{I_m} = \log \frac{\Delta\Omega}{2\pi\tau_0/t_0} + n \log \cos \theta_c. \quad \dots\dots\dots (20)$$

Since, in the Brisbane records, τ_0/t_0 is approximately 0.03, this gives

$$\log I_z/I_m = 3.14 - 0.0257n = 3.14 - 232\theta_c^{-2} \quad (\theta_c \text{ in degrees}).$$

Thus for $\theta_0 = 6^\circ$ the Z -ray intensity should be about 84 dB lower than that of the O -ray.

Experimentally, the Z -ray traces, when appearing on $h'f$ records, have the anticipated narrow width and, on swept-gain records, a horizontal upper edge. The relative intensities of Z - and O -rays have been measured in three cases for which θ_0 was near 6° . These were -75 , -70 , -55 dB. In view of the neglect of absorption, the discrepancy is not serious. It is possible that, in the last case, reflection took place from a tilted part of the ionosphere, thus reducing the angle between the ionization gradient and the magnetic field, and hence the value of θ_c .

At the minimum gain of the swept-gain receiver O -ray echoes are rarely recorded, even when the critical frequency is near 2.28 Mc/s. Since the maximum gain is only 80 dB above this, and since the fixed-gain receivers normally operate with gain intermediate between these two levels, it is not surprising that the Z -ray is rarely recorded.

X. CONCLUSIONS

It is concluded that:

(a) From an examination of the patches produced by a swept-gain fixed-frequency ionospheric recorder, it is possible to gauge the degree of roughness of ionospheric layers.

(b) The effective roughness is a function of the separation of transmitter and receiver, being less the greater the distance between them.

(c) The intensity of Z-ray echoes recorded in Brisbane is consistent with Ellis's theory.

XI. ACKNOWLEDGMENTS

The author wishes to acknowledge his indebtedness to the Radio Research Board of C.S.I.R.O., and to the University Research Committee, whose financial grants rendered this investigation possible, and to Dr. J. A. Thomas for helpful comments.

XII. REFERENCES

- BRIGGS, B. H., and PHILLIPS, G. J. (1950).—*Proc. Phys. Soc. Lond.* B **63**: 907.
BURKE, M. J., and JENKINSON, I. S. (1957).—*Aust. J. Phys.* **10**: 378.
DIEMINGER, W. (1951).—*Proc. Phys. Soc. Lond.* B **64**: 142.
ELLIS, G. R. (1956).—*J. Atmos. Terr. Phys.* **8**: 43.
McNICOL, R. W. E., and WEBSTER, H. C. (1956).—*Aust. J. Phys.* **9**: 272.
McNICOL, R. W. E., WEBSTER, H. C., and BOWMAN, G. G. (1956).—*Aust. J. Phys.* **9**: 247.
MOORE, R. K., and WILLIAMS, C. S. (1957).—*Proc. Inst. Radio Engrs., N.Y.* **45**: 228.
STROHFELDT, M., McNICOL, R. W. E., and GIPPS, G. DE V. (1952).—*Aust. J. Sci. Res.* A **5**: 464.

SOLAR BRIGHTNESS DISTRIBUTION AT A WAVELENGTH OF 60 CENTIMETRES

II. LOCALIZED RADIO BRIGHT REGIONS

By G. SWARUP* and R. PARTHASARATHY†

[Manuscript received March 31, 1958]

Summary

The localized radio bright regions on the Sun which give rise to a slowly varying component of the solar radiation were studied at a wavelength of 60 cm, using a 32-aerial interferometer with a beamwidth of 8.7 min of arc. The observations were undertaken during July 1954 to March 1955 and were limited in number due to this being a minimum period of the solar cycle. The low activity, however, provided the advantage of simple interpretation as often only one region was present on the solar disk.

The characteristics of the observed bright regions are described; their occurrence is closely correlated with regions of chromospheric faculae, the emission polar diagram has a half-power width of 6 days, the estimated size of the sources varied from less than 3 to 6 min of arc, the largest value of the derived brightness temperature was 10^7 °K, and for two groups of localized regions the height of the source was derived to be $35,000 \pm 15,000$ km above the photosphere. Sometimes the slowly varying component showed marked intensity fluctuations in periods of nearly half an hour. The presence of apparently associated fluctuations suggests that at least a part of the slowly varying component at 60 cm has a non-thermal origin.

I. INTRODUCTION

Measurements of solar radiation at decimetre wavelengths by several workers identified a component which has been called the slowly varying component. It varies slowly from day to day, and was distinguished from the rapid variations that were observed occasionally for a duration of seconds or minutes. Pawsey and Yabsley (1949) found it to be superimposed upon a basic steady level attributed to thermal radiation from the "quiet Sun". The component was found by Covington (1948) and Lehany and Yabsley (1949) to have good correlation with the sunspot area. Eclipse observations made by Covington (1947) at a wavelength of 10 cm, and by Christiansen, Yabsley, and Mills (1949) at 50 cm showed that the component originated in localized regions on the solar disk. These radio bright regions were observed close to positions where sunspots were visible or had occurred during previous rotations. One was found close to a stable prominence well off the limb.

To investigate the characteristics of bright regions in detail, it was necessary to isolate them from the background radiation of the quiet Sun by using aerials of

* Division of Radiophysics, C.S.I.R.O., University Grounds, Chippendale, N.S.W.; present address: Stanford University, Stanford, California.

† Division of Radiophysics, C.S.I.R.O., University Grounds, Chippendale, N.S.W.; present address: Radio Propagation Unit, National Physical Laboratory of India, New Delhi.

narrow beamwidth. Christiansen, Warburton, and Davies (1957) studied the regions at a wavelength of 21 cm during the years 1952-53, using a 32-aerial interferometer of beamwidth 3 min of arc. Covington and Broten (1954) and Dodson (1954) made a study at 10 cm using a wave-guide array with a beamwidth of 7.5 min of arc. Kakinuma (1956) made observations at 7.5 cm using an 8-element interferometer with quarter-wavelength plates that had a beamwidth of 4.5 min of arc. It was shown that the occurrence of bright regions on the Sun at 10 and 21 cm is closely correlated with regions of chromospheric faculae, seen in calcium or hydrogen line emission. The study at 7.5 cm confirmed the presence of a small degree of circularly polarized component.

We have now employed Christiansen's 32-aerial interferometer to study the solar radiation at a wavelength of 60 cm. The results concerning the brightness distribution across the quiet Sun have been reported in Part I of this series (Swarup and Parthasarathy 1955). The study of the localized bright regions which give rise to the slowly varying component is described here.

The observations were undertaken during July 1954 to March 1955, which was during the minimum period of the solar cycle. The radio emission from the disturbed Sun was low for most of the period, which is consistent with the low level of optical activity. Only a small number of localized regions that gave strong radio emission were observed. This limited our study of the regions. Because of this, it should be emphasized that the results obtained here are more of an exploratory nature. The low level of activity has had the advantage, however, that often only one bright region was present on the solar disk at any one time. This simplified the interpretation of observations made with an interferometer which had a limited resolving power.

II. OBSERVATIONS

The 32-aerial interferometer described in Part I (Swarup and Parthasarathy 1955) produces a family of fan-shaped beams like those of a diffraction grating. The angular spacing of the beams is 4.9° , which is much larger than the diameter of the solar disk. The half-power beamwidth is 8.7 min of arc. The aerial pattern was derived by measuring the response of the radio source Cygnus-A and that of a strong, localized, bright region on the solar disk. The fan-shaped beams of the interferometer scan the solar disk in a stripwise manner producing a series of records of the one-dimensional radio brightness distribution across the solar disk each day. The daily records exhibit peaks which change in position with solar rotation. This shows the presence of localized bright regions on the solar disk. As described in Part I, the response curves of the localized bright regions for any day are obtained by subtracting the background component due to the quiet Sun from the record of the one-dimensional distribution for the day. The derived curves give both the position and intensity of the radio bright regions. The location is given along a line corresponding to the position of the fan-shaped aerial beam on the solar disk when the peak appears on the record.

Taking the flux density of the quiet Sun as $2 \times 10^{-21} \text{ W m}^{-2} (\text{c/s})^{-1}$ on two polarizations, the error in derivation of flux densities of bright regions was estimated to be $3 \times 10^{-23} \text{ W m}^{-2} (\text{c/s})^{-1}$. The error was small, as the calibration

of records was checked in terms of power received from adjacent parts of the quiet Sun. Receiver fluctuations were negligible. The total effective area of the 32 parabolic aerials was nearly 35 m^2 , and the measured transmission line loss was 3 dB. The receiver had a noise factor of 10 dB, bandwidth of 4 Mc/s, and time constant of 2 sec.

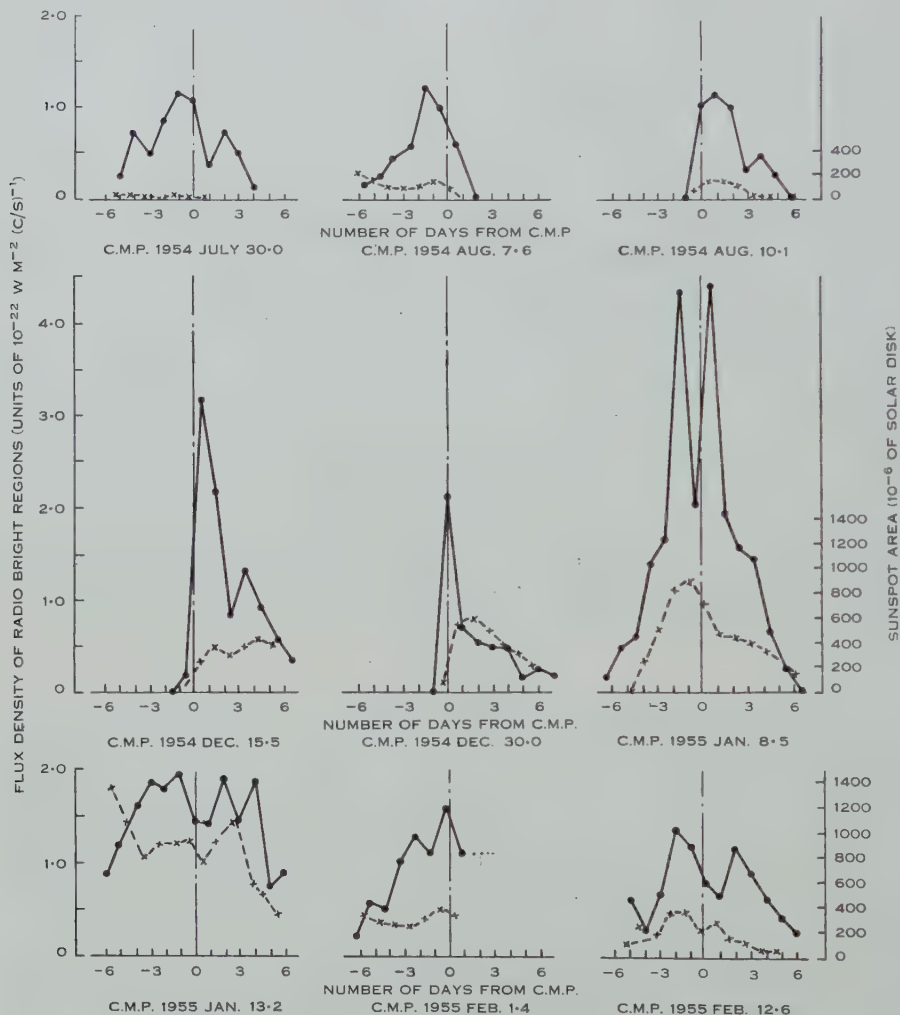


Fig. 1.—Intensity of radio bright regions plotted against time separation, in days, from central meridian passage (full line). Area of the associated sunspots is also plotted (broken line).

III. OBSERVED CHARACTERISTICS OF RADIO BRIGHT REGIONS

(a) Emission Polar Diagram

During the period July 1954 to March 1955 several isolated radio bright regions were studied during their passage across the solar disk. Figure 1 shows the radio flux densities for nine of the most active regions observed during the period. The projected area of the associated sunspots is also plotted in the figure for comparison.

It is clear from Figure 1 that, apart from variations due to growth or decay of a radio bright region, its radiation flux density is greater when it is located closer to the central meridian of the Sun. An attempt was, therefore, made to calculate the directivity or the emission polar diagram of an average radio bright region at a wavelength of 60 cm. It was assumed that the emission polar diagram was symmetrical with respect to the central meridian, and that for the nine regions observed any variations due to growth or decay occurred randomly with respect to their location on the solar disk. The latter assumption was considered to be justified on an examination of the associated sunspot areas. Only those observa-

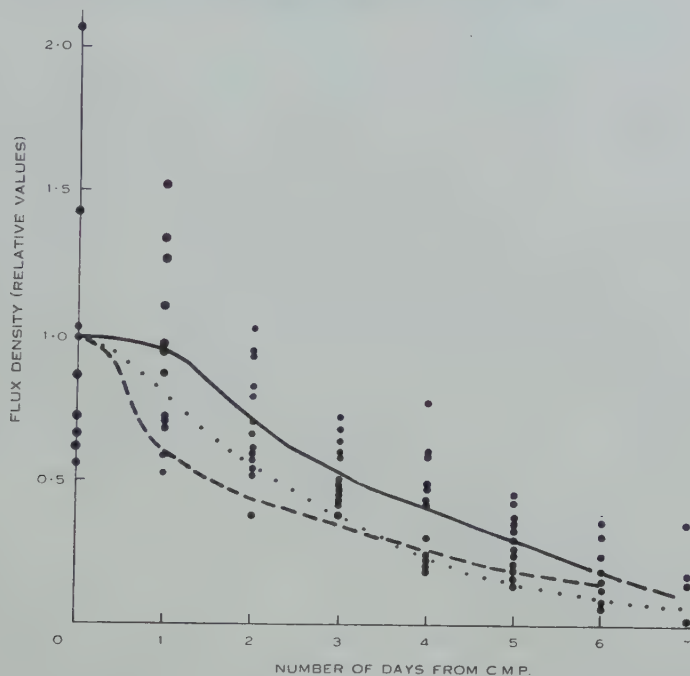


Fig. 2.—Emission polar diagram derived by us (full line), by Machin and O'Brien (broken line), and by Müller (dotted line). The values of flux density measured by us, after being normalized, are shown by dots.

tions were considered for this study for which the radio bright regions were seen near their central meridian passage (C.M.P.) and for the adjoining 5 days on the eastern or western side of the solar disk. A mean was taken of the flux densities of each region measured for the above 6 days. The mean values were used to normalize the curves shown in Figure 1. The normalized values are plotted in Figure 2 against the time separation, in days, from C.M.P. of the regions. The values of days are approximated to the nearest integer. The curve passing through the mean of the points shown in Figure 2 is the derived emission polar diagram. The variation with position for the radio regions at 60 cm is steeper than the cosine variation known to occur for the visible area of sunspots.

Machin and O'Brien (1954) calculated the emission polar diagram for a wavelength of 60 cm by a statistical analysis of the daily apparent disk temperature

of the Sun measured during 1953. Müller (1956) deduced the polar diagram at 50 cm from a study of the apparent disk temperature during the passage of eight large sunspot groups in the years 1949-1952. The curves derived by them are also shown in Figure 2 and are reasonably similar to ours. However, we would like to point out again that the present observations, though limited in number, had small errors and were free from confusion as the radio bright regions were isolated from the background.

(b) Source Height above Photosphere

As mentioned earlier, the interferometer gives the position of a radio bright region on the solar disk along a line. Its location, however, can be determined unambiguously by a method described by Christiansen and Warburton (1953) if it is observed for several days during its rotation on the solar disk. On a Mercator's projection of the Sun, an arc is drawn for each day corresponding to the heliographic coordinates of the line that passes through the bright region on the solar disk. The arcs intersect at a point only if the bright region lies on the photosphere. Several similar graphs are drawn, assuming different values of effective solar radius corresponding to distance of the source from centre of the Sun. The best intersection gives heliographic coordinates as well as height of the bright region. The method has only one assumption, that the heliographic coordinates of the region remain the same from day to day; this implies a rotational velocity of the region of 13.2° per day. However, a correction can be made for a different rotational velocity of the region, if it is known.

During December 1954 and January 1955 two bright regions were observed on the solar disk over several days and gave sufficiently strong radio emission for their positions to be located within $\pm \frac{1}{4}$ min of arc. The mean rotational velocity of the associated sunspot groups was 12.5° per day for each group. For derivation of height, it was assumed that the bright regions had the same rotational velocity. The height above the photosphere was determined to be $35,000 \pm 15,000$ km for both the regions. There were two main sources of error, uncertainty in measurements of the positions of the radio source, and change of heliographic coordinates of the source in an uncertain manner due to its proper motion or due to any absorption effects taking place as the source approached the limb.

(c) Size and Equivalent Brightness Temperature

For an extended source the response pattern of the aerial is wider than that for a point source. The approximate size of the source can be deduced from this widening if an assumption is made as to its shape. We assumed the bright regions on the solar disk to be circularly uniform. The only cases analysed were those where the radio bright regions were completely isolated and there was only a single area of optical activity within the region of response of the aerial beam.

The radio bright regions observed at 60 cm produced only a small amount of broadening. This was reliably determined only when the sources were sufficiently strong, and the study was limited to these. The largest calculated size was 6 min of arc, which was obtained for the radio source observed on January 15 and 17, 1955. Some strong sources did not show any detectable widening, for which an upper limit to their diameters was calculated to be 3 min of arc.

The equivalent temperature of a bright region was calculated using the measured value of its flux density and its estimated size. The largest value obtained was 10^7 °K which occurred on December 16, 1954 at 0144 U.T. It should be noted that the radio source, which was probably not uniform, was not fully resolved and the peak value of brightness temperature could be higher. The deduced value of size, average daily flux density, and the corresponding equivalent temperature for the various days are listed in Table 1. Values of the associated

TABLE 1
SIZE AND EQUIVALENT BRIGHTNESS TEMPERATURE OF RADIO BRIGHT REGIONS

Date	C.M.P. of the Region	Projected Sunspot Area (10^{-6} disk units)	Diameter (min of arc)	Average Flux (10^{-22} W m $^{-2}$ (c/s) $^{-1}$)	Brightness Temperature (10^6 °K)
1954, Dec. 16	Dec. 15.5	240	<3	3.2	>7.4
Dec. 17	Dec. 15.5	310	<3	2.2	>5.1
Dec. 19	Dec. 15.5	415	<3	1.3	>2.1
Dec. 30	Dec. 30.0	240	<3	2.1	>4.9
1955, Jan. 5	Jan. 8.5	260	<3	1.4	>3.1
Jan. 14	Jan. 13.2	830	5	1.4	1.2
Jan. 15	Jan. 13.2	820	6	1.9	1.1
Jan. 17	Jan. 13.2	530	6	1.9	1.1
Jan. 18	Jan. 13.2	415	4	0.8	1.0
Jan. 19	Jan. 13.2	300	4	0.9	1.1
Average		440	<4	1.7	>2.8

sunspot areas are also given in the table for comparison. It is interesting to note that the values obtained for the size and equivalent temperature of the sources at 60 cm are similar to those derived at 50 cm by Christiansen, Yabsley, and Mills (1949) from observation of an eclipse in 1948. The total sunspot area was 850×10^{-6} of the solar disk on the day of the eclipse. Eight sources were identified whose estimated sizes varied between 2 and 4 min of arc. The average equivalent temperature was 5×10^6 °K, the most intense being at least 10^7 °K.

(d) *Association with the Optically Active Regions on the Sun*

The bright regions observed at a wavelength of 60 cm were compared with the optical features on the Sun. It was found that radio bright regions were always associated with regions of chromospheric faculae or plages. Except for some weak regions, sunspots appeared in these active regions for at least some part of their lifetime. The radio bright regions often preceded sunspots by a few days and succeeded them for a much longer period. The activity of bright regions usually increased with the appearance of sunspots.

The relationship between the flux densities of radio bright regions and the areas of the associated sunspots can be investigated using the curves shown in Figure 1. It has been shown in Section III (a) that the observed radio emission

decreases as the region rotates from the centre to the limb. Also the visible sunspot area has a cosine variation. It has been pointed out by Christiansen, Warburton, and Davies (1957) that this common dependence on position might give a fictitiously high correlation between radio emission and sunspot area. The dependence on position was, therefore, eliminated for the radio regions by using suitable correction factors obtained from the emission polar diagram shown in Figure 2. Correspondingly, in the case of sunspots, values of projected area were used. As it was undesirable for the correction factors to become too large, only

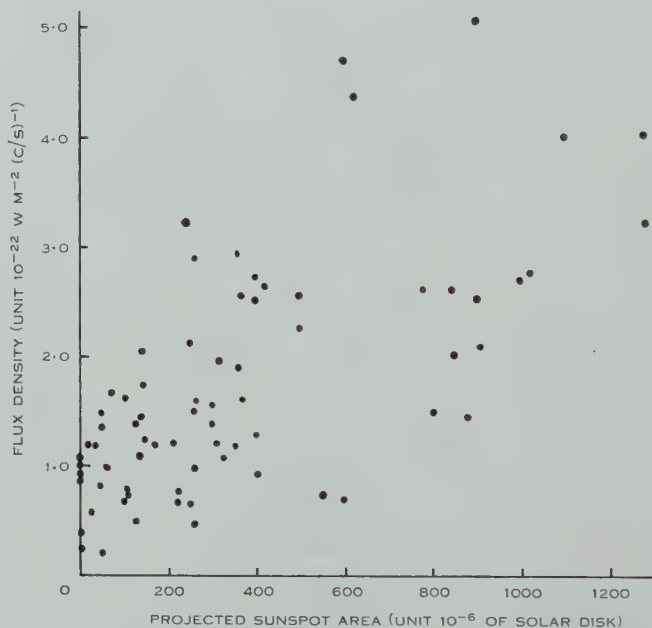


Fig. 3.—Scatter diagram showing flux of a radio bright region versus projected sunspot area. The radio flux has been corrected for directivity of the source (see Fig. 2).

those regions were considered which were situated within 5 days of the C.M.P. The corrected values of radio emission are plotted in a scatter diagram, Figure 3, against projected area of the associated sunspots. The correlation coefficient between the two was found to be 0.62.

(e) *Variation of Brightness in a Short Period*

Previous observations of the total radio emission from the Sun at decimetre wavelengths have shown that the slowly varying component remains noticeably steady for a period of a few hours and usually varies slowly from day to day. Our measurements, however, occasionally gave evidence of marked fluctuations in the brightness of the localized regions which gave rise to the slowly varying component, in periods of half an hour.

Records of the one-dimensional distribution across the solar disk were obtained by the interferometer over a period of about 2 hr around midday. Five

records separated by nearly 20 min were obtained in this period as the successive interferometer beams scanned the solar disk. Any variations in the brightness of a localized region could be detected quite accurately in the successive records as any gain change of the receiver could be checked by comparing in the two

TABLE 2
FLUX DENSITIES OF FLUCTUATING BRIGHT REGIONS
($10^{-22} \text{ W m}^{-2} (\text{c/s})^{-1}$)

Date	Relative Time				
	0 ^m	21.5 ^m	43 ^m	1 ^h 4.5 ^m	1 ^h 26 ^m
1954, Dec. 16	3.0	3.0	3.0	4.2	2.8
Dec. 19	0.7	0.7	1.5	1.9	1.5
Dec. 30	2.3	2.0	2.4	1.9	1.9
1955, Jan. 7	4.0	4.6	4.8	4.1	4.1
Jan. 15	1.6	1.6	1.8	2.6	
Jan. 17	2.2		2.2	1.6	1.6

records power received from adjacent parts of the quiet Sun. From December 15, 1954 to January 19, 1955, when four large sunspot groups moved across the solar disk, the flux density of the associated radio bright regions was more than 5 per cent of the quiet Sun flux density for 22 cases; in such cases it was thought that

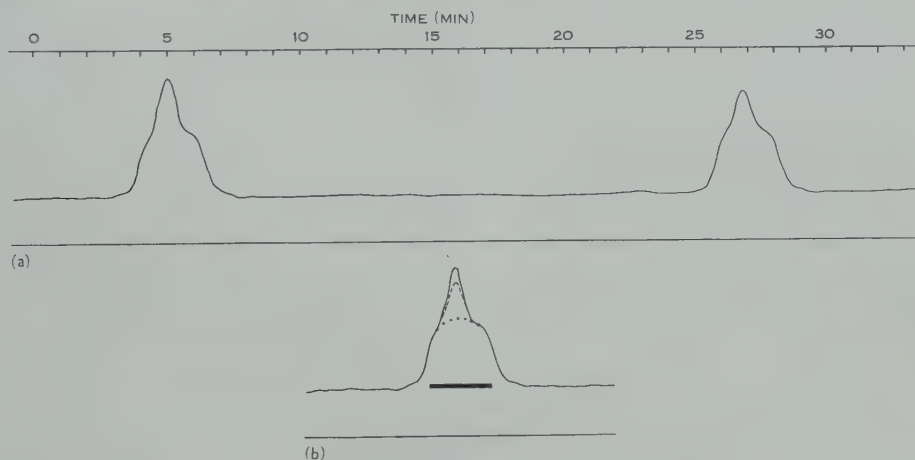


Fig. 4 (a).—Two successive records obtained on December 16, 1954.

Fig. 4 (b).—The two records shown in (a) are superimposed upon each other. The background radiation is shown by the dotted line and the size of the visible solar disk by the thick line.

a relative change of 15 per cent or more in brightness of the regions was detectable, being outside the limits of experimental errors. Variations were noted only in five cases for which the values of flux density for the five patterns are given in Table 2. Two successive records obtained on December 16, 1954, are shown in Figure 4 (a). The records are superimposed in Figure 4 (b), which illustrates the

observed change. Though the brightness of the localized regions changed, the shape of the response curves obtained on subtracting the background component remained the same in all five cases. This implies that the variations are slow, as it takes about 1·25 min for the interferometer beams to scan the bright regions. The changes are slower than that occurring at the time when bursts are observed in solar radiation. Moreover, no bursts were observed on the above occasions at frequencies of 200 and 600 Mc/s. Only in one case (January 15, 1955) a noise storm occurred at a frequency of 200 Mc/s. At two other times when a noise storm occurred at a frequency of 200 Mc/s no detectable variation was observed in our records at 500 Mc/s. The changes observed in the present observations are considered to have been caused by comparatively rapid variations of the slowly varying component.

It can be seen from Table 2 that on two occasions a variation of more than 40 per cent. and on one occasion a variation of more than 100 per cent. in the brightness have been noted in a period of nearly 20 min.

IV. COMPARISON WITH CHARACTERISTICS DERIVED AT OTHER WAVELENGTHS

Some of the characteristics of radio bright regions determined by various high resolution observations at decimetre wavelengths are summarized in Table 3. The differences in the characteristics with respect to the wavelength are in the proper direction to fit qualitatively in the existing ideas about the radio bright regions.

The emission polar diagram of the radio source becomes narrower as the wavelength increases. This result is the same as derived by Müller (1956) and others from statistical analysis of observations of the whole Sun. However, the emission polar diagram calculated from observations of the whole Sun is narrower for each wavelength than that derived from high resolution observations.

It should be noted that the heights of the radio source at the three wavelengths were derived for different bright regions. Moreover, experimental errors were large. It is therefore difficult to make a comparison. It is likely that the height of the radio source above the photosphere increases with the wavelength, although the possibility of origin from nearly the same level cannot be excluded.

Observations at decimetre wavelengths show that the size of the radio bright regions corresponds approximately to that of the chromospheric faculae (Ca or H α plages). Also, Dodson (1954) found good correlation between intensities of the radio regions at 10 cm with those of calcium plages. Recently it was shown by Hatanaka *et al.* (1955a) that the brightness isophotes of a radio region at 8 cm agreed well in shape with a calcium plage. All these observations suggest that there is a very close relation between the origins of the radio emission and line emission from chromospheric faculae.

We compared our 60 cm observations with the 7·5 cm observations (simultaneous in time) undertaken by Kakinuma (1956), who employed an 8-element interferometer with beam of 4·5 min of arc. It was found that the ratio of the radio flux at 7·5 cm to that at 60 cm, when the regions were near

C.M.P., varied between 3 and 10 for different regions. If we assume nearly the same size of the source, it is seen that the brightness temperature of the radio region is appreciably higher at the longer wavelength. The values of the bright-

TABLE 3
CHARACTERISTICS OF BRIGHT REGIONS DETERMINED BY HIGH RESOLUTION OBSERVATIONS AT DM-CM WAVELENGTHS

Technique	Multiple-element Interferometer Observations			Eclipse Observations	
Period of observations	1954	1952-53	1955	Nov. 1, 1948	June 20, 1955
Observers	Authors	Christiansen, Warburton, and Davies (1957)	Kakinuma (1956)	Christiansen, Yabsley, and Mills (1949)	Hatanaka <i>et al.</i> (1955 <i>b</i>)
Characteristics	Wavelength (cm)				
	60	21	7.5	50	8
1. Emission polar diagram	Steeper than cosine variation	Cosine variation			
Half-power width ..	6 days	9 days			
2. Height above photosphere (km)	35,000 $\pm 15,000$	24,000 $\pm 3,000$	35,000		
3. Observed sizes (min. of arc)	<3 to 6	<3 to 10	3 to 5	2 to 4	4
4. Brightness temperature of a strong region (10^6 °K)	5	2	0.7	5	1
5. Flux density near C.M.P. for an associated sunspot area of 10^{-3} of the Solar disk (in units of 10^{-22} W m $^{-2}$ (c/s) $^{-1}$)	3.5	10	18*		
Flux density of quiet Sun (10^{-22} W m $^{-2}$ (c/s) $^{-1}$)	20	30	80		

* This information was evaluated by the authors from the drift curves of the interferometer published by the observers (*Proc. Res. Inst. Atmosph., Nagoya University* (1955) 3: 149).

ness temperature for the radio regions that are given in Table 3 agree broadly with the hypothesis of thermal origin of the radio emission (Waldmeier and Müller 1950; Piddington and Minnett 1951).

It was shown by Christiansen, Warburton, and Davies (1957) that at 21 cm the slowly varying emission corresponding to the same sunspot area decreases to nearly half from the maximum (1947) to the minimum (1952-53) of a solar cycle,

and so does the quiet Sun flux. The study made at 60 cm gives a similar indication though the effect is less pronounced. At 60 cm both of the components decreased by a factor of 0.7 approximately.

V. CONCLUSIONS

The observations, made with an aerial of high resolving power operating at a wavelength of 60 cm, enabled us to study several highly emitting regions on the solar disk which give rise to the slowly varying component. It was found that the radio regions always occurred in association with Ca or H α plages. It was possible to detect radio emission even from a very faint plage region. Except for some weak radio regions, sunspots appeared in these for some part of their lifetime. A correlation coefficient of 0.62 was found between the flux from the radio regions and the area of the sunspots.

The study gave information about the emission polar diagram, the height above the photosphere, the size, and the equivalent brightness temperature of the regions at 60 cm. The results are summarized in Table 3. A comparison was made with the results derived at shorter decimetre wavelengths by different workers. The characteristics change with wavelength in the proper direction to fit qualitatively the existing ideas about the origin of radio emission from these active regions on the solar disk. The simplest explanation is that the radio emission originates in thermal radiation from regions of very high electron density and temperature such as "coronal condensations" (Waldmeier and Müller 1950) which tend to occur in the corona over sunspot groups. In order to make quantitative deductions it is desirable to make simultaneous optical and radio observations of the same region.

Christiansen, Warburton, and Davies (1957) pointed out that no evidence appeared against the hypothesis of thermal origin in their investigation at 21 cm. However, the present observations give some evidence that at least part of the slowly varying component at 60 cm has a non-thermal origin. On six occasions, there were marked fluctuations in the brightness of the radio regions in periods of half an hour. It is difficult to visualize these fluctuations on the hypothesis of thermal origin as it takes time to change the temperature of a large volume of gas (70,000 km in extent). As a further indication of the non-thermal origin, it should be noted that, whereas radio brightness temperatures over 10^7 °K have been measured, such high optical temperatures have not.

VI. ACKNOWLEDGMENTS

The authors are indebted to Mr. C. Fryar for assistance in the observational work; and to Dr. J. L. Pawsey, Dr. W. N. Christiansen, and Mr. J. A. Warburton for valuable discussions throughout the period of the project. During the period of observations both the authors were holders of Fellowships awarded by the Australian Government under the Colombo Plan.

VII. REFERENCES

- CHRISTIENSEN, W. N., and Warburton, J. A. (1953).—*Aust. J. Phys.* **6**: 190.
CHRISTIENSEN, W. N., Warburton, J. A., and DAVIES, R. D. (1957).—*Aust. J. Phys.* **10**: 491.
CHRISTIENSEN, W. N., YABSLEY, D. E., and MILLS, B. Y. (1949).—*Aust. J. Sci. Res. A* **2**: 506.

- COVINGTON, A. E. (1947).—*Nature* **159**: 405.
- COVINGTON, A. E. (1948).—*Proc. Inst. Radio Engrs., N.Y.* **36**: 454.
- COVINGTON, A. E., and BROTEN, N. W. (1954).—*Astrophys. J.* **119**: 569.
- DODSON, H. W. (1954).—*Astrophys. J.* **119**: 564.
- HATANAKA, T., AKABANE, K., MORIYAMA, F., TANAKA, H., and KAKINUMA, T. (1955a).—*Publ. Astr. Soc. Japan* **7**: 161.
- HATANAKA, T., AKABANE, K., MORIYAMA, F., TANAKA, H., and KAKINUMA, T. (1955b).—*Rep. Ionosph. Res. Japan* **9**: 195.
- KAKINUMA, T. (1956).—*Proc. Res. Inst. Atmos. Nagoya University* **4**: 78.
- LEHANY, F. J., and YABSLEY, D. E. (1949).—*Aust. J. Sci. Res. A* **2**: 48.
- MACHIN, K. E., and O'BRIEN, P. A. (1954).—*Phil. Mag.* **45**: 973.
- MÜLLER, H. (1956).—*Z. Astrophys.* **39**: 160.
- PAWSEY, J. L., and YABSLEY, D. E. (1949).—*Aust. J. Sci. Res. A* **2**: 198.
- PIDDINGTON, J. H., and MINNETT, H. C. (1951).—*Aust. J. Sci. Res. A* **4**: 131.
- SWARUP, G., and PARTHASARATHY, R. (1955).—*Aust. J. Phys.* **8**: 487.
- WALDMEIER, M., and MÜLLER, H. (1950).—*Z. Astrophys.* **27**: 58.

FLARE-PUFFS AS A CAUSE OF TYPE III RADIO BURSTS

By R. G. GIOVANELLI*

[*Manuscript received April 17, 1958*]

Summary

An analysis has been made of the association of type III solar radio bursts with flares which show sudden bright expansions, or "puffs". Type III bursts occur within ± 2 min of the times of two-thirds of the puffs; by comparison, they occur during the whole lifetimes of about one-quarter of flares of all types (Loughhead, Roberts, and McCabe 1957). Most if not all puffs are followed by surges, and hence it appears that the same explosion yields two quite different ejections; one, at velocities of the order of one-fifth that of light, is responsible for the burst, and the other, at velocities of the order of 100 km/sec, is the surge.

I. INTRODUCTION

In a solar radio burst of type III (Wild, Roberts, and Murray 1954) the frequency of maximum intensity decreases rapidly in a few seconds, the accepted cause being a disturbance ejected at velocities of the order of one-fifth that of light through the solar corona.

Loughhead, Roberts, and McCabe (1957) found that 60-70 per cent. of type III bursts occur during lifetimes of flares of apparent area (i.e. uncorrected for any foreshortening) exceeding about 20×10^{-6} of the Sun's hemisphere. However, rather less than one flare in four is associated with a type III burst. This low figure indicates that there are differences of degree or type between various flares; as over two-thirds of the type III bursts are associated with tiny flares of apparent area below 40×10^{-6} of the Sun's hemisphere, flare size alone is not the determining factor.

It is well known that some flares are associated with other types of ejection. In particular, surges dark in $H\alpha$ sometimes appear on the disk during or after flares. These have outward velocities of the order of only 100 km/sec, and cannot themselves be the disturbances directly responsible for type III bursts, though the chances of a surge accompanying a flare are about the same as those of a type III burst.

Recently, Giovanelli and McCabe (1958) have shown that the first trace of a surge is usually the ejection from the flare of diffuse matter, brighter than the chromosphere, which fades and disappears by becoming transparent, though it can be detected if the ejection crosses the limb. The dark surge subsequently appears close to the flare and travels outwards in the same direction. The basic feature of the event is the emission of a substantially continuous particle stream having different appearances at different stages.

* Division of Physics, C.S.I.R.O., University Grounds, Chippendale, N.S.W.

Close examination of films of flare-surges shows that some of the flares, but not all, appear to be of an explosive character. Within a minute or so the flare may expand rapidly (e.g. Giovanelli and McCabe 1958, Plate 1, Figures 1 and 2; Plate 2, Figure 2; Plate 3, Figure 1). Usually, but not always, this expansion or "puff" occurs at the outbreak of the flare. However, the initial bright expansion is much more diffuse and indefinite in some of the other flare-surges (e.g. Giovanelli and McCabe 1958, Plate 2, Figure 1), and these are not included here as "puffs".

Its violent and sudden nature suggests that a puff is the identifiable flare characteristic most likely to be associated with type III bursts. The following account describes an investigation to establish whether or not this is so. It is found that there is a very strong and significant association, and that the times of type III bursts and puffs are closely identical.

II. OBSERVATIONS AND ANALYSIS OF RESULTS

The flare observations were made with the Sydney H α flare-patrol camera, 16 mm diameter photographs of the Sun being obtained on Eastman Kodak IV E film, usually at 0.5 min intervals. These have been examined in projection. Type III burst data obtained with the Dapto radiospectrograph have been provided by Mr. J. P. Wild.

The analysis covers observations made in 1956-1957, from which there has been prepared a list of flares showing clear, indisputable puffs occurring during the times of radiospectrograph observation. The list is small (27 flares) and incomplete, but the events are reasonably well distributed over the 2-year period. Times of puffs have then been compared with times of type III bursts.

TABLE 1
ASSOCIATION OF TYPE III BURSTS WITH FLARE PUFFS

Apparent area (10^{-6} hemisphere)				0-40	41-80	> 80	Total
Number of flare puffs				14	6	7	27
Puffs with co-incident type III bursts	(a) Total			11	4	3	18
	(b) Cases with non-puff flares beginning at puff time ± 4 min			2	2	1	5

Results are summarized in Table 1, which gives (i) the numbers of flare-puffs grouped according to apparent flare area, (ii) the number of these with a type III burst occurring within ± 2 min of the time of the puff, and (iii) the number of the latter group in which puffs occurred within ± 4 min of the beginning of another flare. With one exception, all the above flares produced dark surges.

Table 1 shows that 18 of 27 puffs, or 67 per cent., coincided with type III bursts to within ± 2 min. Most of the flares were of class 1 —, 20 having apparent areas below 80×10^{-6} of the Sun's hemisphere. The apparent decrease in frequency of association as the flare area increases is not statistically significant.

The analysis of Loughhead, Roberts, and McCabe has shown that type III bursts occur during the lifetimes (± 2 min) of only about 22 per cent. of all flares. There is therefore negligible chance of the close association found between puffs and type III bursts being fortuitous. Loughhead, Roberts, and McCabe have also shown that in 50 per cent. of associations the type III bursts occur within ± 2 min of the beginning of the flare. In five of the present cases non-puff flares began within the rather larger limits of ± 4 min of the time of the puff. At most, these would be expected to produce only about one association, insufficient to affect the conclusions.

The present investigation establishes that there is a close and significant association between flare-puffs and type III bursts, two-thirds of the flare-puffs producing bursts within ± 2 min of the time of the puff. No significant dependence on flare size has been established.

III. DISCUSSION

The question naturally arises whether the flare-puff is the only flare phenomenon which produces type III bursts. It is not possible to answer this as yet, for the small size of the majority of flares makes it very difficult to distinguish puffs on the small-scale flare-patrol photographs. However, it is of interest that Loughhead, Roberts, and McCabe (1957) reported an investigation in which 48 flare-surges were studied, type III bursts being found to occur during 29 per cent. of them, whereas such bursts occurred during 18 per cent. of 246 flares not accompanied by observed surges. Their data were insufficient to establish that flare-surges are more likely to be associated with type III bursts than flares not ejecting surges. But, from the relatively low rate of association with flare-surges as a whole, it appears that type III bursts are much more likely to be associated with flare-surge events showing rapid expansions (puffs) than with those where the initial bright expansion is more diffuse and indefinite.

The present conclusions add appreciably to our knowledge of the flare-surge event. As shown by Giovanelli and McCabe, puffs, in common with other flare expansions, are the first stages in the ejection of particle streams with velocities of the order of 100 km/sec; in $H\alpha$, these streams are bright, transparent, or dark against the Sun's disk, depending on excitation conditions which vary with time along the stream. Within about a minute or so, i.e. effectively simultaneously, the same explosions eject at velocities of the order of one-fifth that of light the high speed disturbances responsible for type III bursts.

IV. ACKNOWLEDGMENTS

The flare observations described here were made under the supervision of Miss Marie K. McCabe; details of burst observations have been provided by Mr. J. P. Wild.

V. REFERENCES

- GIOVANELLI, R. G., and MCCABE, MARIE K. (1958).—*Aust. J. Phys.* **11**: 191.
LOUGHHEAD, R. E., ROBERTS, J. A., and MCCABE, MARIE K. (1957).—*Aust. J. Phys.* **10**: 483.
WILD, J. P., ROBERTS, J. A., and MURRAY, J. D. (1954).—*Nature* **173**: 532.

OPTICAL OBSERVATIONS OF THE SOLAR DISTURBANCES CAUSING TYPE II RADIO BURSTS

By R. G. GIOVANELLI* and J. A. ROBERTS†

[Manuscript received May 12, 1958]

Summary

Identifications have been established for the solar optical events associated with a number of type II radio bursts. Near or at the limb these have been ejections with velocities exceeding that of sound in the corona. Where the event has been on the disk there has usually been a very bright flare, with some evidence of dark filament activity. In two cases the event was the disappearance (i.e. ejection) of a filament with the subsequent development of flares on either side.

Evidence is presented that the fundamental plasma frequency is emitted from events much further from the centre of the disk than expected on the basis of a spherically symmetrical corona.

I. INTRODUCTION

Solar radio bursts of spectral type II are characterized by a steady decrease in frequency over a period of a few minutes (Wild 1950). This has generally been interpreted as due to the passage of an ejected particle stream through the corona at a velocity of the order of 500 km/sec, exciting plasma oscillations whose frequency is proportional to $N_e^{\frac{1}{2}}$. Type II bursts are relatively infrequent, and, while it is known that they are associated with chromospheric flares, there has been no account of direct observation of the particle streams causing them.

In many events a fundamental band and its second harmonic can both be observed. On the plasma hypothesis it is expected that in the case of a spherically symmetrical corona this will happen only when the source is close to the centre of the disk; for sources beyond about 20° the fundamental band should not escape (Wild, Murray, and Rowe 1954).

The investigation described below has been aimed not only at identifying the particle streams responsible for type II bursts but also at establishing whether, well away from the centre of the disk, emission is confined to the second harmonic of the plasma frequency.

II. OBSERVATIONS

The radio-frequency records have been obtained with the Dapto radio-spectrograph (40–240 Mc/s), the optical records with the Sydney $H\alpha$ flare-patrol camera.

During the period 1956–1957, 18 type II bursts were recorded during times of optical observation. In 13 of these cases we believe that we have been able

* Division of Physics, C.S.I.R.O., University Grounds, Chippendale, N.S.W.

† Division of Radiophysics, C.S.I.R.O., University Grounds, Chippendale, N.S.W.

TABLE I
LIST OF TYPE II BURSTS AND IDENTIFICATIONS

Date	Type II		Optical Event					
	Time	Harmonics Observed	Type	Time	Coordinates	Radial ⁽¹⁾ Distance	Prominence Velocity ⁽²⁾ (km/sec)	Uncorrected ⁽²⁾ Flare Area
1. iii.57	0035-0059	—	Ejected prominence (L) ⁽³⁾	0010-0032	21° S., 90° + W.	1.00 +	$\begin{cases} 350^{(4)} \\ 280^{(5)} \\ 490^{(6)} \end{cases}$	—
26. iii.57	0412-0416	—	Ejected prominence (L)	0357-0411	11° N., 90° + E.	1.00 +		—
24. ix.57	0212-0226	—	Flare-surge (L)	0200-0225	28° S., 90° W.	1.00	$\begin{cases} \geq 200 \\ \sim 670 \end{cases}$	10
24. ix.57	0504-0508	—	Surge (L)	0453-0456	28° S., 90° W.	1.00		—
31. i.57	0407-0424	1, 2	Disappearing filament followed by two parallel flares	0358-0550	25° N., 5° E.	0.50	1050	1050
29. xi.57	0101-0103	—	Disappearing filament followed by two parallel flares (L)	0038-0600	41° N., 63° E.	0.95		600
6.viii.57	0431-0438	1, 2	Diffusely expanding flare	0423-0433	25° N., 22° W.	0.43		70

⁽¹⁾ In projected solar radii.

⁽²⁾ $\times 10^{-6}$ Sun's hemisphere; uncorrected for any foreshortening.

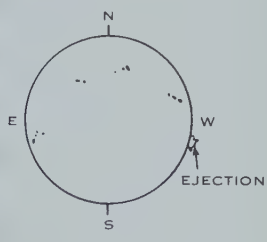
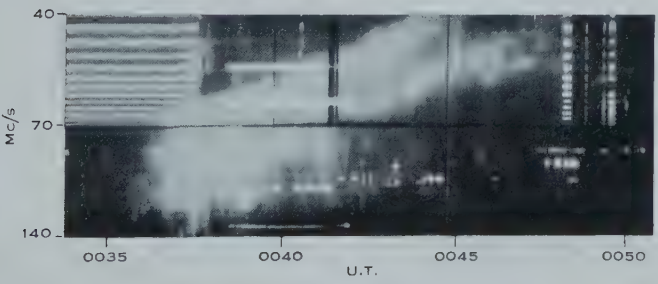
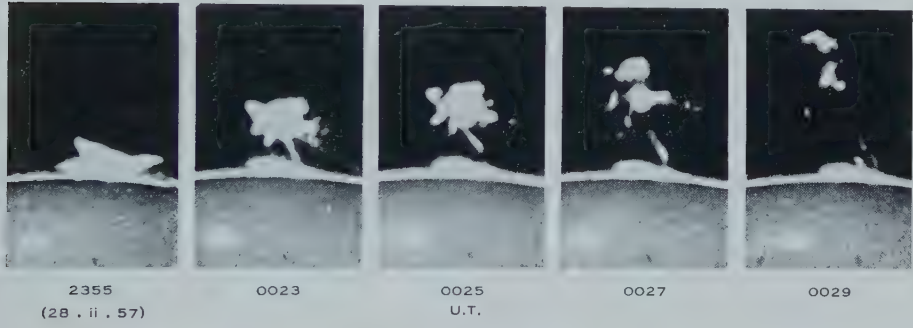
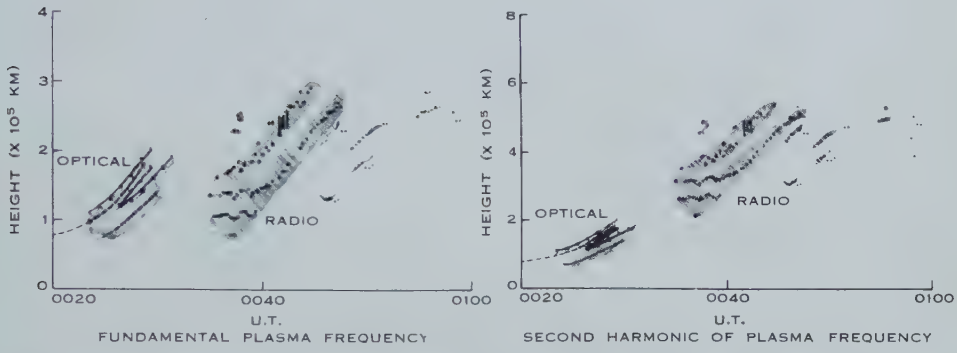
⁽³⁾ Letter (L) indicates ejection observed beyond limb.

⁽⁴⁾ Velocity of top of ejection.

⁽⁵⁾ Velocity of top of lower part of ejection.

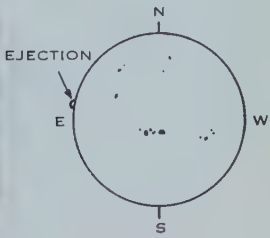
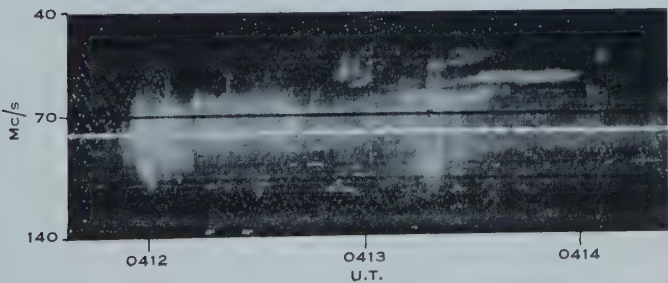
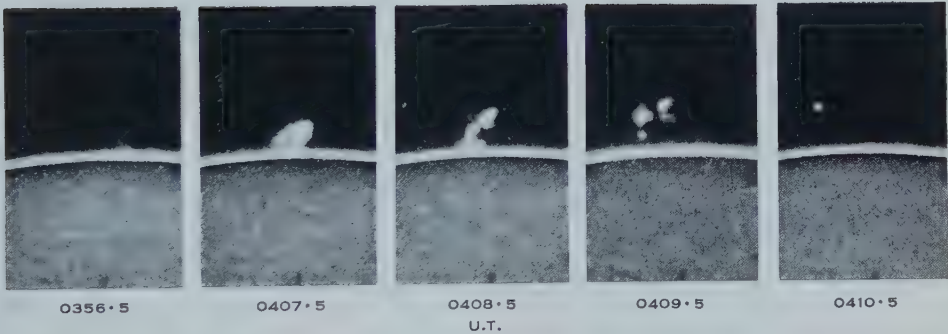
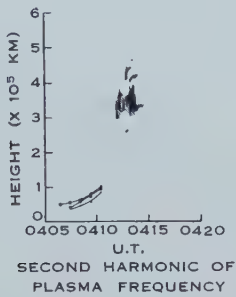
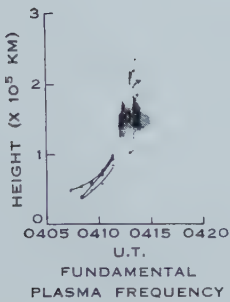
⁽⁶⁾ Velocity of fastest knots.

OPTICAL AND RADIO SOLAR DISTURBANCES



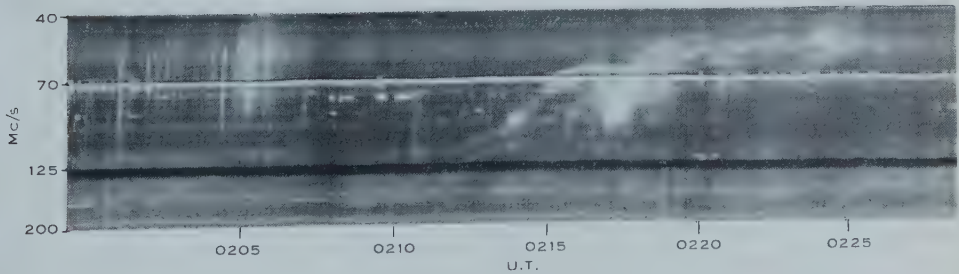
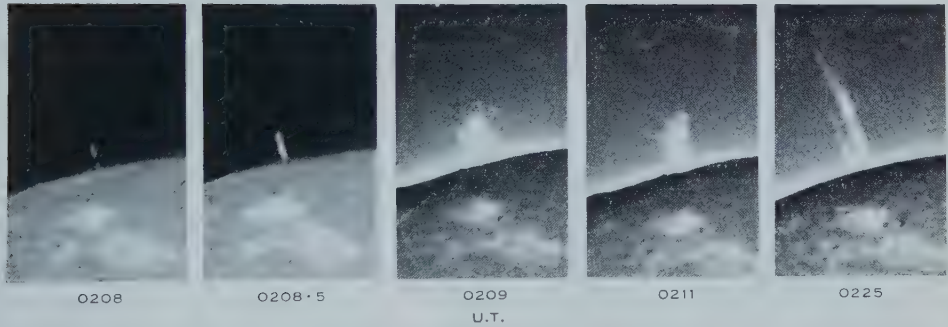
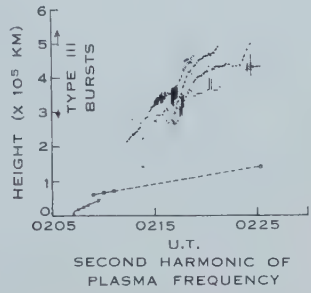
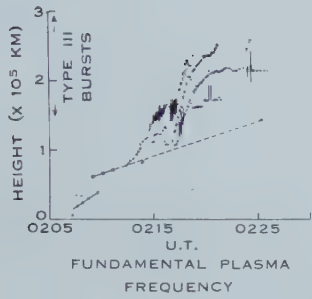
MARCH 1, 1957

OPTICAL AND RADIO SOLAR DISTURBANCES



MARCH 26, 1957

OPTICAL AND RADIO SOLAR DISTURBANCES



SEPTEMBER 24, 1957

OPTICAL AND RADIO SOLAR DISTURBANCES

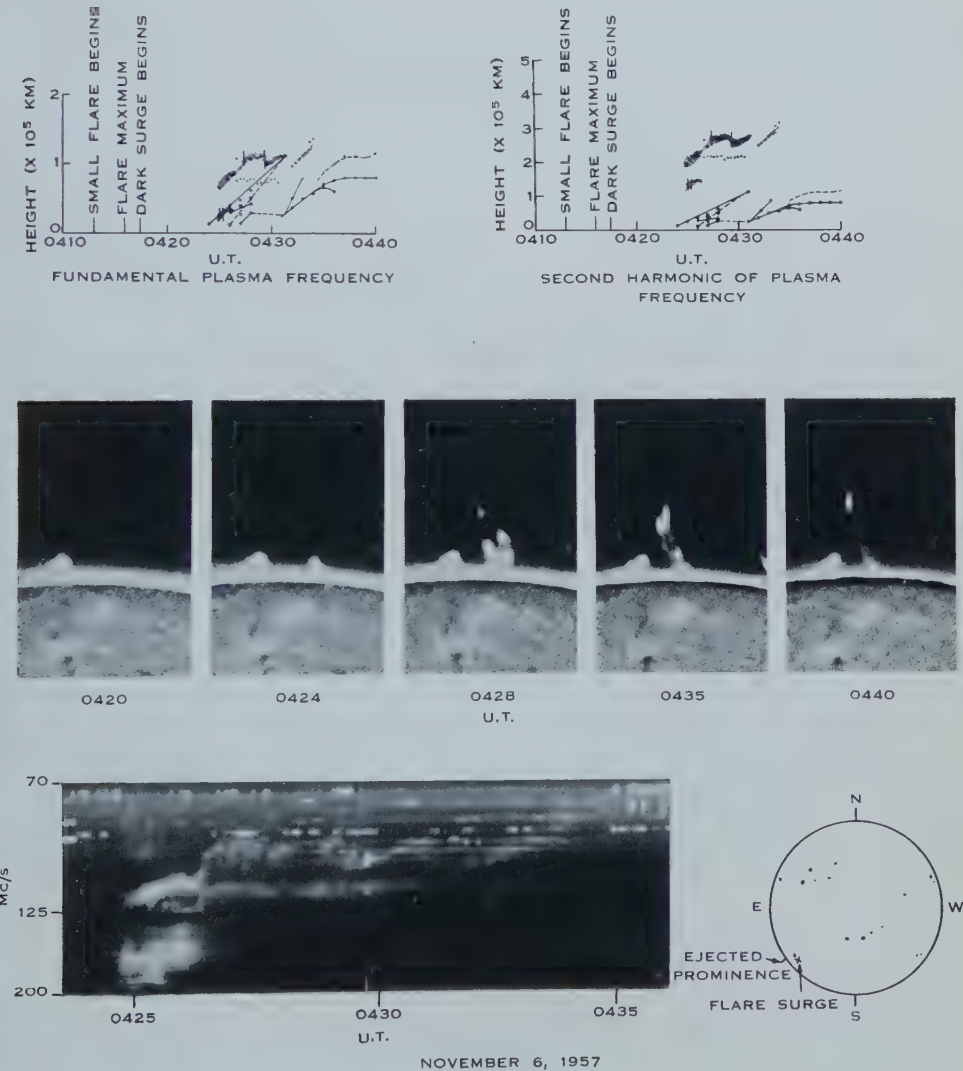


TABLE 1 (Continued)

Date	Type II		Optical Event					Un- cor- rected Flare Area
	Time	Harmonics Observed	Type	Time	Coordinates	Radial Distance	Prominence Velocity (km/sec)	
19. i.56	0026-0031	1, 2	Brilliant flare causing dis- appearance of filament	0016-0046	22° N., 26° E.	0.56		450
13. xi.56	0207-0220	Inadequate record	Flare with a short-lived dark marking appearing on disk	0141-0306	23° N., 47° W.	0.77		385
13. xi.57	0502-0505	1, 2	Brilliant flare	0457-0511	25° S., 27° E.	0.59		170
22. xi.57	0411-0426	1, 2	Bright flare showing extensions and an associated dark marking	0404-0440	32° N., 26° W.	0.61		385
6. xii.57	0400-0419	1, 2	Very bright flare, causing dis- appearance of nearby filament	0345-0440	17° N., 44° E.	0.75		350
20. xii.57	0547-0552	1, 2	Brilliant flare with an associated dark marking	0550 ⁽⁷⁾	17° N., 2° W.	0.34		175
17. vii.57	0125-0132	1, 2	Flare ⁽⁸⁾ Disappearing filament ⁽⁸⁾	0112-0142 0113-0128	12° N., 31° E. 10° N., 50° E.	0.48 0.80		220 —
6. xi.57	0425-0434	—	Ejected prominence (L) ⁽⁸⁾ Flare-surge ⁽⁸⁾	0424-0445 0413-0450	38° S., 90° + E. 33° S., 55° E.	1.00 + 0.90	240	— 30

⁽⁷⁾ Incomplete observation.⁽⁸⁾ Uncertain identification.

to identify the optical disturbance responsible for the bursts. In two additional cases there are alternative identifications. In the remaining three cases the optical records are poor or interrupted. A list of the identified events in Table 1 gives the classification of the optical disturbance, date, times of optical disturbance and of type II bursts, disk coordinates, radial distance, and harmonics observed.

Four of the events originated at or beyond the limb, and in a fifth case (November 6, 1957) one of the alternative identifications is with an ejected limb prominence. Ejected matter was also observed beyond the limb in the event of November 29, 1957, during which a filament on the disk disappeared and two parallel lines of flares developed on either side of its former position. Of the limb events, those of March 1 and March 26, 1957 were well observed. The record was interrupted during the earlier event of September 24, 1957; the later event on the same day was too confused for detailed analysis. Fairly good records were obtained of the limb event of November 6, 1957.

Plates 1-4 show selected photographs of the limb events on March 1, March 26, the earlier event on September 24 and on November 6, 1957, together with radiospectrograph records of the same events. On the same plates, graphs are given of the height-time plots obtained from the optical and radio observations. The latter have been derived by assuming that the emission occurs at the coronal plasma frequency or its second harmonic, and that the distribution of electron density with height follows the Allen-Baumbach model. There is no clear harmonic structure in any of these cases, and, as it is not obvious whether the observed emission is in the fundamental or second harmonic band, two alternative graphs are given.

III. DISCUSSION OF RESULTS

Examination of Table 1 shows that type II bursts have been identified with a range of optical events. Of the 13 definite identifications, 9 are with flares on the disk, 2 with surges at the limb, and 2 with ejected prominences at the limb. Seven of the flares on the disk occurred in the neighbourhood of sunspots; almost all were particularly bright and were active in the sense of ejecting surges or showing evidence of dark hydrogen activity. Two flares followed the disappearance of filaments, a sequence studied in detail by Bruzek (1957); such events involve the ejection of prominences, as was in fact observed at the limb in one of our cases (November 29, 1957) and on the disk in the other (January 31, 1957).

The linear velocities of all these events were high, in excess of 200 km/sec. This is greater than the velocity of sound, which is 170 km/sec in ionized hydrogen at a temperature of 10^6 °K. It may be noted that on September 24, 1957, several other surges not accompanied by type II bursts were observed in the region responsible for the two bursts listed in Table 1. The velocities in these events were lower, of the order of 160 km/sec or less.

The above evidence indicates that type II bursts are mostly associated with bright flares showing particle ejections. It is then of interest to examine the limb events illustrated in Plates 1-4 to ascertain whether the particle ejections observed in $H\alpha$ are the exciting agencies responsible for the type II bursts.

Comparison of the height-time plots of the optical and radio events on March 1, 1957 (Plate 1) suggests that the type II burst is the second harmonic emission of coronal plasma oscillations excited by the ejected prominence material. The appearance of the second harmonic rather than the fundamental is in accordance with theory. The correlation between the optical and radio events is quite remarkable when one considers that the radio heights are based on averaged electron densities from a number of eclipse observations.

In the other cases (Plates 2, 3, and 4) the agreement is not nearly so good, but favours the emission of radiation at the fundamental plasma frequency. The second event on September 24, 1957 (not illustrated) favours the second harmonic.

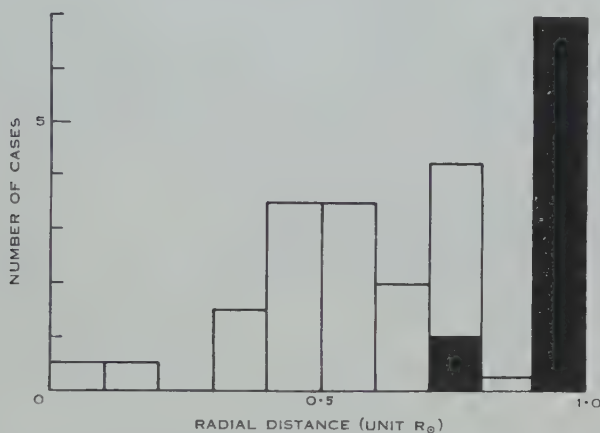


Fig. 1.—Distribution of type II bursts with radial distance. Shaded regions refer to bursts not showing harmonic structure, while unshaded regions refer to bursts with fundamental and second harmonic bands.

From the study of these five limb events, we conclude that the ejections seen in $H\alpha$ could well be the exciting agency responsible for type II bursts. However, the case is by no means proven. If the emission at the limb occurs only in the second harmonic, this demands some additional assumption such as (a) the extension of the ejected particle stream beyond the observed ejected prominence; longer exposure times could possibly test this in future events; or (b) a lower electron density than given by the Allen-Baumbach formula; or (c) departures from spherical symmetry. The alternative (c) is not unexpected, since the corona shows a strong structure of bright streamers and dark lanes, and therefore possesses considerable density variations.

The present results may be used in a different way to determine whether the fundamental band is emitted from the limb. In Figure 1 we give a histogram showing the distribution of the radial distances of the bases of the optical events associated with the type II bursts. This figure also includes additional probable identifications using flares recorded elsewhere. The part of this histogram which is shaded refers to bursts that show no evidence of harmonic structure,

the unshaded portion applies to bursts showing both fundamental and second harmonic bands. The two harmonics almost always appear when the radial distance is $\lesssim 0.75$, but are rarely observed together in events nearer the limb. These results confirm theory to the extent of showing that limb events are observed only in one band (presumably the second harmonic). However, the fundamental band is observed considerably further from the centre of the disk than expected on the spherically symmetric model.

Clearly, some modified theory is required taking account of coronal structure, while more extensive observations are needed of limb ejections, with direct measurements of the position and movement of the radio sources.

IV. ACKNOWLEDGMENTS

The optical observations used in this study were made under the supervision of Miss Marie McCabe, who also assisted with the analysis. The radio observations were made by Messrs. K. V. Sheridan, G. H. Trent, and J. Joisce, under the direction of Mr. J. P. Wild.

V. REFERENCES

- BRUZEK, A. (1957).—*Z. Astrophys.* **42**: 76.
WILD, J. P. (1950).—*Aust. J. Sci. Res. A* **3**: 399.
WILD, J. P., MURRAY, J. D., and ROWE, W. C. (1954).—*Aust. J. Phys.* **7**: 439.

EXPLANATION OF PLATES 1-4

Each of the Plates 1-4 illustrates one of the limb events and contains the following four sets of data.

(1) Height-time plots derived from the optical and radio observations. Optical results are shown by points joined by full or broken lines, and the radio results by a series of dots and shading. The two height-time diagrams given for the radio event have been derived on the assumption of emission at (a) the fundamental plasma frequency of the corona, or (b) the second harmonic of the plasma frequency. The Allen-Baumbach distribution of electron density has been assumed.

(2) A series of photographs of the optical event, selected to show its development.

(3) The dynamic radio spectrum, printed with an inverted frequency scale so as to simulate a height-time plot. Sharp features in these spectra, and in particular continuous or intermittent lines at discrete frequencies, are caused by interfering or calibrating signals.

(4) A sketch of the solar disk showing the location of the optical event.

PLATE 1

Event of March 1, 1957. The two arms of the height-time diagram for the optical event (shaded in this figure) refer to the two parts of the ejected prominence which are visible in the photographs at 0027 and 0029 U.T.

PLATE 2

Event of March 26, 1957. Identifiable knots in the ejected prominence were used to form the height-time diagram. Note that the time scale of the spectrograph record is very different from those in the other plates.

PLATE 3

Event of September 24, 1957. The first two photographs of the flare surge show the development of the flare. Enhanced limb photographs shortly afterwards show the flare to be surrounded by a cloud of fainter matter in the process of ejection. The record was interrupted between 0211 and 0225 U.T., during which time there had clearly been a surge. The radio spectrograph record shows a group of type III bursts coinciding closely with the beginning of the flare.

PLATE 4

Event of November 6, 1957. There are alternative identifications for the optical event corresponding to this type II burst. The photographs show clearly an ejected limb prominence, for which the heights of various features are plotted in the height-time diagrams. Not so clearly visible on the photographs is a faint dark surge ejected towards the limb from a tiny flare in contact with the upper left of the sunspot at the bottom of the photographs. The flare had almost subsided by 0420.

A CATALOGUE OF RADIO SOURCES BETWEEN DECLINATIONS $+10^\circ$ AND -20°

By B. Y. MILLS,* O. B. SLEE,* and E. R. HILL*

[Manuscript received March 3, 1958]

Summary

A catalogue has been prepared of the radio sources observed between declinations $+10^\circ$ and -20° , using the Sydney cross-type radio telescope at a wavelength of 3.5 m: a total of 1159 sources is listed in the area of 3.24 steradians. This supersedes an earlier catalogue of Mills and Slee in portion of the area, but the differences between the two are small. A number of new identifications with galaxies are suggested, and an analysis made of the statistics of the source distribution. It is concluded that cosmological effects displayed by the distribution, if present, are small. Possibilities are discussed of separating from this distribution the effects of the instrument, the finite angular sizes, and/or the physical clustering of the sources.

I. INTRODUCTION

A preliminary catalogue of radio sources has been published in the declination interval $+10^\circ$ to -20° and between Right Ascensions 00^h and 08^h (Mills and Slee 1957). The observations were made with the Sydney cross-type radio telescope at a wavelength of 3.5 m. This catalogue was preliminary in the sense that adequate checking had not been possible in some places, but was published at that time to draw attention to the large discrepancies between it and the catalogue of Shakeshaft *et al.* (1955) of Cambridge.

The area concerned has now been checked thoroughly and extended between the same declination limits over the whole 24 hr of Right Ascension; in the additional area a great deal more observational data have been used than previously. Other catalogues of more southerly regions are in preparation and will be published in due course.

Very few drastic changes have been needed in the preliminary catalogue, but a large number of minor improvements were possible as the result of accumulating more observations. Accordingly, it has been found desirable to repeat in full the list of sources in the original area. A total of 1159 sources is now listed between $+10^\circ$ and -20° ; the area involved is 3.24 steradians.

These more extensive data strengthen the conclusions reached in the earlier paper. There is again extremely poor detailed agreement with the Cambridge catalogue; moreover, the new statistics agree closely with our earlier more limited results but disagree grossly with the Cambridge statistics. It is concluded, as before, that the source count ogive is not significantly different from that expected with a uniform spatial distribution of sources. However, there is

* Division of Radiophysics, C.S.I.R.O., University Grounds, Chippendale, N.S.W.

evidence for a slightly excessive slope which could be due to a *small* excess of faint sources, physical clustering of the sources, or simply a statistical deficiency of stronger sources in our immediate neighbourhood. These possibilities are considered in Section IV. The catalogue has been examined for identifications with various celestial bodies using the Skalnate Pleso Catalogue (Becvar 1951) and some parts of the Palomar Sky Atlas that are now available. There are a number of possible new identifications with extragalactic nebulae, including the Hydra II and Pegasus I and II clusters of galaxies: these are very briefly discussed.

II. THE CATALOGUE

The preparation of the catalogue, given in Tables 1, 2, and 3, follows the methods outlined by Mills and Slee (1957); that paper will henceforth be referred to as paper I. Over most of the area, at least two normal "scanning" type records were available (i.e. quasi-simultaneous recordings on five declinations); in addition, "non-scanning" records had been taken in a large number of areas. In general, sources were systematically picked out using the best of the available scanning records and checked against all other records of the same region. During the checking process, a few additional weak sources were noted which had been obscured on the original records by noise fluctuations or interference; these were also included. There remain a few areas where checking was not possible, either because of the absence of a second record or because of interference. Sources in these areas are indicated in the catalogue and must generally be regarded as less reliable than the others; experience indicates that the main effects of using a single record are to miss some of the weaker sources and to introduce rather large errors in the flux densities. In estimating the positions and flux densities of the remaining sources the best possible use of all the data has been made by taking averages or, where indicated, weighted averages of the measurements available. This procedure has resulted in slight modifications to at least one of the three listed measurements of a substantial proportion of the sources in paper I. Flux densities, in general, are given to two significant figures, but for the weaker sources the second figure has no significance and has been retained only for consistency in the statistics. Regions where possible confusion arises with bright sources in the side lobes have been indicated by Mills *et al.* (1958).

The same general method of cataloguing has been adopted as in paper I, and, for convenience, a reference number has been attached to each source which is defined in a similar way. The first two digits of the reference number denote the hour of the Right Ascension; these are followed by the sign and tens digit of the declination in degrees and an italicized serial number arranged in order of increasing Right Ascension within the 1-hr period. To save space, however, only the italicized serial numbers are given in the catalogue as the others are immediately evident; for example in Table 1, the second source listed would be referred to in the text as 00+02.

As before, the probable errors in the final digit of a position measurement have been estimated and are indicated by superscripts. Some comparisons

TABLE 1
SOURCES BETWEEN DECLINATIONS $+10^\circ$ AND 0°

Sources observed on one record only are indicated by an asterisk. Sources which may be "extended", that is, resolvable, are indicated by a dagger. A colon has been placed beside uncertain flux densities

Ref. No.	Position (1950)		Flux Density (10^{-26} $\text{W m}^{-2} (\text{c/s})^{-1}$)	Ref. No.	Position (1950)		Flux Density (10^{-26} $\text{W m}^{-2} (\text{c/s})^{-1}$)
	R.A. h m	Dec. N. ° '			R.A. h m	Dec. N. ° '	
	00				02		
1	04.9 ³	06 05 ⁸	35 (20) ⁽¹⁾	4	12.5 ⁵	06 11 ⁸	14
2	10.2 ²	00 37 ⁵	20	5	19.4 ²	08 08 ⁶	24
3	14.2 ³	06 48 ⁷	18	6	26.2 ³	02 35 ⁷	9.0
4	16.0 ⁴	08 20 ⁸	11	7	35.7 ³	07 01 ⁸	10
5	24.8 ⁴	07 28 ⁶	11	8	50.4 ⁴	01 19 ⁶	10
6	30.0 ⁶	01 40 ¹⁰	68 (20) ⁽²⁾	9	53.4 ³	06 48 ⁷	11
7	30.8 ³	05 53 ⁵	25	10	55.1 ¹	05 53 ⁴	51
8	32.1 ⁴	04 28 ⁶	16 ⁽³⁾	11	58.9 ³	01 35 ⁵	27
9	34.2 ³	00 12 ⁸	15				
10	36.7 ⁴	03 35 ⁶	14		03		
11	37.4 ²	09 30 ⁵	37	1	00.2 ³	07 20 ⁶	18
12	40.1 ³	06 53 ⁷	14	2	00.9 ³	09 37 ⁶	13
13	40.6 ⁴	02 20 ⁸	7:	3	05.4 ²	03 50 ⁵	34
14	42.6 ⁴	05 26 ⁷	19	4	09.2 ⁴	05 11 ⁸	10
15	55.3 ³	01 14 ⁶	16	5	25.2 ³	02 25 ⁵	41
16	55.5 ⁴	08 47 ⁸	14	6	34.1 ⁴	09 51 ⁷	24
17	59.8 ³	04 32 ⁵	19	7	35.8 ⁴	07 40 ⁷	13
				8	40.5 ³	04 55 ⁵	35
				9	45.6 ³	00 41 ⁵	15
				10	46.6 ³	05 42 ⁶	15
				11	51.4 ⁴	02 58 ⁶	16
				12	58.2 ²	00 27 ⁵	19
	01				04		
1	14.9 ³	06 05 ⁶	11	1	00.0 ³	05 35 ⁸	13
2	16.3 ²	08 11 ⁶	14	2	00.1 ³	02 21 ⁵	11
3	17.3 ⁴	03 20 ⁵	33 (16) ⁽⁴⁾	3	04.7 ²	03 45 ⁴	37
4	23.1 ³	01 22 ⁵	20 ⁽⁵⁾	4	11.9 ⁴	05 43 ⁷	8.6
5	24.4 ³	09 13 ⁷	16	5	21.9 ⁴	00 24 ⁸	14
6	28.7 ²	03 52 ⁶	18	6	23.2 ⁴	04 26 ⁶	13
7	29.2 ²	06 07 ⁵	23	7	28.6 ²	01 02 ⁵	20†
8	33.5 ⁴	07 53 ⁶	15	8	32.8 ²	03 57 ⁵	25†
9	34.9 ⁴	06 34 ⁸	7.4	9	38.2 ⁵	07 05 ⁸	8: ⁽⁷⁾
10	43.0 ³	02 01 ⁶	9.4	10	41.8 ³	02 15 ⁵	30†
11	46.2 ³	06 10 ⁸	19†	11	51.5 ⁴	02 33 ⁶	10
12	47.2 ³	07 07 ⁶	19†	12	54.9 ⁵	06 43 ⁸	13
13	52.1 ⁴	03 32 ⁶	49 (27) ⁽⁶⁾	13	56.3 ³	05 20 ⁸	10
14	57.4 ³	01 10 ⁶	16	14	58.5 ⁴	01 24 ⁶	15
	02						
1	02.3 ⁴	04 20 ⁸	10				
2	07.4 ²	09 35 ⁸	23				
3	11.0 ²	02 58 ⁵	24				

(1) Perhaps two sources.

(2) Complex brightness distribution; may be several sources.

(3) Perhaps a background irregularity. (4) (NGC 470/474). (5) (NGC 533).

(6) Perhaps two sources or interference from 05N2A.

(7) Doubtful, not visible on all records.

TABLE 1 (Continued)

Ref. No.	Position (1950)		Flux Density (10^{-26} $\text{W m}^{-2} (\text{c/s})^{-1}$)	Ref. No.	Position (1950)		Flux Density (10^{-26} $\text{W m}^{-2} (\text{c/s})^{-1}$)
	R.A. h m	Dec. N. ° ' "			R.A. h m	Dec. N. ° ' "	
	05				09		
1	04.5 ³	07 20 ⁷	16	1	09.2 ³	08 23 ⁷	13
2	10.9 ³	01 02 ⁶	38	2	15.2 ²	09 35 ⁶	40 ⁽¹³⁾
3	16.4 ⁴	09 58 ⁸	20	3	34.1 ²	04 50 ⁵	24
4	16.5 ²	03 59 ⁸	17	4	34.4 ⁴	02 13 ⁶	14†
5	28.9 ³	06 35 ⁶	30	5	41.8 ³	09 57 ⁷	32
6	38.8 ⁴	05 43 ⁸	9.2	6	43.0 ³	02 21 ⁵	7.1
7	41.5 ³	02 46 ⁶	22† ⁽⁸⁾	7	44.9 ²	07 39 ⁴	89
				8	49.7 ²	00 06 ⁵	36
				9	50.5 ³	09 00 ⁷	16
				10	55.3 ²	03 35 ⁵	18
	06				10		
1	00.5 ⁴	02 23 ⁷	11	1	05.7 ²	07 54 ⁶	30
2	02.3 ³	00 54 ⁵	12	2	08.6 ³	06 32 ⁶	39
3	05.4 ⁵	08 08 ¹⁰	109 (28)	3	09.9 ⁴	04 50 ¹⁰	8.7
4	14.2 ⁴	05 43 ⁸	18†	4	10.9 ³	03 11 ⁶	9.4
5	15.3 ⁴	03 36 ⁸	8.8	5	22.0 ³	09 36 ⁸	15
6	20.3 ⁴	09 00 ¹⁰	153 (45)	6	24.0 ³	06 41 ⁶	35*
7	24.8 ³	02 50 ⁵	18*	7	38.2 ³	02 36 ⁸	13
8	29.6 ²	05 01 ³	250 (87) ⁽⁹⁾	8	47.3 ³	04 25 ⁷	13†
9	32.6 ³	02 09 ⁴	29	9	48.8 ²	00 00 ⁵	21†
10	34.8 ³	07 15 ⁸	72 (40)	10	54.0 ⁴	02 09 ⁶	24 (14) ⁽¹⁴⁾
11	42.7 ⁴	05 15 ⁸	27	11	56.8 ⁴	09 15 ⁷	19
12	42.8 ⁴	00 10 ¹⁰	16				
13	52.5 ⁴	03 00 ⁸	22				
14	54.1 ³	08 36 ¹⁰	24†				
	07				11		
1	17.9 ⁴	08 48 ⁸	11	1	06.5 ⁵	09 45 ⁷	16†
2	19.4 ³	01 34 ⁵	17* ⁽¹⁰⁾	2	07.1 ⁴	03 48 ⁷	15
3	29.6 ³	03 06 ⁸	21	3	08.0 ⁵	01 56 ⁸	7.0
4	41.9 ²	02 05 ⁵	36	4	20.2 ⁴	07 40 ⁵	8.2*
5	44.9 ⁴	09 57 ⁸	13	5	20.9 ³	05 25 ⁷	19
6	53.7 ⁴	07 10 ⁷	8.2 ⁽¹⁰⁾	6	22.5 ⁴	02 20 ⁸	8.5
				7	26.3 ⁴	00 42 ⁸	7.1
				8	37.4 ³	01 24 ⁷	15
	08			9	38.4 ⁴	05 43 ⁸	8.2*
1	03.4 ⁴	04 48 ⁸	8.7	10	42.4 ³	08 14 ⁷	14
2	12.4 ³	01 33 ⁶	29 (22)	11	42.5 ⁵	09 30 ¹⁰	8.6
3	19.8 ²	06 07 ⁴	125 (60) ⁽¹¹⁾	12	47.0 ³	05 40 ⁷	11*
4	33.4 ³	00 42 ⁷	17	13	54.1 ³	04 26 ¹⁰	12†
5	34.5 ³	09 30 ⁷	13	14	59.6 ³	00 36 ⁷	8.5
6	38.5 ⁵	03 17 ⁸	17†				
7	41.0 ⁴	07 28 ⁷	14				
8	43.3 ⁴	02 20 ⁸	9.4				
9	54.9 ⁴	09 55 ⁸	14				
10	55.0 ⁴	03 36 ⁸	9.0 ⁽¹²⁾				

⁽⁸⁾ Could be associated with Barnard's ring. ⁽⁹⁾ (NGC 2237).⁽¹⁰⁾ A doubtful source. ⁽¹¹⁾ Possibly interference from 08S4A.⁽¹²⁾ (Hydra II cluster). ⁽¹³⁾ Possibly interference from 09S1A. ⁽¹⁴⁾ Perhaps two sources.

TABLE 1 (Continued)

Ref. No.	Position (1950)		Flux Density (10^{-26} $\text{W m}^{-2} (\text{c/s})^{-1}$)	Ref. No.	Position (1950)		Flux Density (10^{-26} $\text{W m}^{-2} (\text{c/s})^{-1}$)
	R.A. h m	Dec. N. ° ' "			R.A. h m	Dec. N. ° ' "	
	12				14		
1	01.7 ⁵	07 14 ⁸	13	11	35.5 ³	00 23 ⁶	14
2	04.2 ²	04 19 ⁵	25	12	37.1 ⁵	08 58 ³	12*
3	07.4 ³	08 39 ¹⁰	12	13	40.6 ²	05 04 ⁶	15
4	14.8 ²	04 00 ⁶	30 ⁽¹⁵⁾	14	45.0 ²	07 54 ⁶	19*
5	16.7 ¹	05 59 ⁵	100 ⁽¹⁶⁾	15	56.5 ³	04 01 ⁶	11
6	18.0 ³	09 50 ¹⁰	24				
7	19.0 ²	02 46 ⁶	12*		15		
8	26.6 ²	02 17 ⁴	167	1	00.1 ⁴	06 15 ¹⁰	15*
9	35.3 ³	01 42 ⁸	16	2	08.2 ³	08 09 ⁷	42†
10	46.9 ⁵	09 23 ⁸	14 ⁽¹⁷⁾	3	08.6 ²	06 08 ⁶	24*
11	51.5 ³	08 53 ⁶	17 ⁽¹⁸⁾	4	09.8 ⁵	01 42 ⁸	20
				5	14.2 ²	07 11 ⁶	140
				6	14.4 ⁵	00 18 ⁸	16
1	02.0 ²	09 02 ⁷	22	7	19.3 ²	07 55 ⁶	50*†
2	04.5 ⁴	07 02 ⁷	18	8	33.3 ⁵	09 29 ⁸	13
3	08.0 ³	06 10 ⁷	22*†	9	34.1 ²	02 38 ⁶	17
4	09.7 ²	04 00 ⁷	7.5	10	36.1 ⁴	01 42 ⁷	12
5	12.6 ⁵	07 41 ⁸	16	11	37.4 ³	06 08 ⁸	22*
6	18.7 ³	01 00 ⁸	50 (27)	12	42.1 ³	04 06 ⁷	17*
7	30.3 ³	02 18 ⁸	19	13	42.4 ²	02 20 ⁵	16
8	32.8 ³	06 22 ¹⁰	8.0	14	48.9 ²	03 10 ⁶	18*
9	40.3 ⁴	02 20 ⁹	13				
10	45.6 ³	00 42 ⁶	8.3		16		
11	50.0 ⁵	06 19 ⁷	54 (22)*	1	00.0 ²	02 13 ⁵	100 ⁽²¹⁾
12	55.0 ²	01 24 ⁶	19	2	02.8 ⁴	01 05 ⁷	45
13	55.4 ³	04 50 ⁷	10	3	03.3 ²	00 06 ⁶	35
				4	07.0 ⁴	04 25 ⁷	17
				5	13.3 ³	04 25 ⁶	27
1	01.0 ³	09 22 ⁷	28	6	22.2 ³	08 21 ⁶	14
2	09.4 ³	07 31 ⁶	14	7	29.0 ⁵	09 08 ¹⁰	22*†(22)
3	13.0 ⁵	05 49 ⁸	17	8	38.1 ³	03 44 ¹⁰	23*
4	15.8 ³	01 06 ⁷	13	9	44.7 ³	01 43 ⁸	33†
5	16.7 ¹	06 43 ⁴	114	10	48.8 ¹	05 04 ²	390 ⁽²³⁾
6	17.0 ²	04 00 ⁵	22 ⁽¹⁹⁾				
7	24.3 ³	04 19 ⁷	13*		17		
8	25.4 ⁴	00 36 ⁶	16	1	03.2 ³	09 16 ¹⁰	78 (49)* ⁽²²⁾
9	32.5 ³	06 38 ⁷	17* ⁽²⁰⁾	2	22.3 ³	05 44 ⁴	36*† ⁽²⁴⁾
10	35.5 ²	03 36 ⁶	47 (31)	3	56.1 ⁴	02 46 ¹⁰	48.*

(15) (NGC 4234). (16) (NGC 4261), (NGC 4270).

(17), (18) Perhaps one extended source. (19) (NGC 5566).

(20) Perhaps E.-W. side lobe of Virgo-A.

(21) Perhaps slightly extended.

(22) A doubtful source.

(23) IAU 16N0A.

(24) Perhaps a galactic irregularity.

TABLE 1 (Continued)

Ref. No.	Position (1950)		Flux Density (10^{-26} $\text{W m}^{-2} (\text{c/s})^{-1}$)	Ref. No.	Position (1950)		Flux Density (10^{-26} $\text{W m}^{-2} (\text{c/s})^{-1}$)
	R.A. h m	Dec. N. ° ' "			R.A. h m	Dec. N. ° ' "	
	18				21		
1	03.9 ⁴	00 12 ⁷	33†	4	26.4 ²	07 15 ⁷	27
2	04.3 ⁴	03 40 ¹⁰	27	5	27.1 ⁵	01 06 ⁸	67 (25) ⁽²⁸⁾
3	15.3 ⁴	00 04 ¹⁰	14	6	36.0 ⁴	03 47 ⁶	12
4	17.4 ³	03 05 ⁶	41	7	39.8 ⁴	02 45 ⁷	18†
5	26.5 ⁴	00 25 ⁸	50 (24)	8	42.4 ³	07 54 ⁸	15
6	29.6 ⁴	09 40 ¹⁰	45†	9	42.5 ⁴	04 00 ⁵	11
7	34.6 ³	03 37 ⁶	20:†	10	49.6 ⁴	07 52 ⁸	23
8	42.7 ⁷	09 30 ¹⁰	54†	11	50.5 ⁴	05 17 ⁷	17
9	43.3 ⁴	07 15 ⁸	28:	12	52.2 ⁴	02 08 ⁷	15
10	44.0 ⁵	05 07 ⁸	25†	13	58.3 ⁵	05 16 ⁷	8.3*
11	53.7 ¹	01 29 ⁵	550 ⁽²⁵⁾	14	59.9 ³	04 25 ⁵	8.4
	19				22		
1	09.0 ³	05 05 ⁸	59† ⁽²⁶⁾	1	06.5 ⁴	01 54 ⁷	26
2	12.7 ³	00 09 ⁸	29†	2	10.0 ⁵	07 41 ⁶	14
3	17.5 ³	00 54 ¹⁰	20†	3	10.8 ⁵	08 48 ⁷	31*†
4	30.1 ³	00 54 ⁷	20	4	21.4 ⁵	02 17 ⁷	8.1
5	32.4 ⁴	09 43 ⁶	30:	5	22.4 ⁴	05 55 ⁵	16
6	33.8 ³	05 55 ⁸	18† ⁽²⁶⁾	6	26.6 ³	08 27 ⁷	19
7	37.4 ³	04 13 ¹²	34†	7	34.9 ⁴	05 43 ⁸	12
8	37.7 ⁴	01 06 ¹⁰	8.3	8	39.9 ⁴	04 28 ⁸	12
9	43.8 ³	09 16 ⁸	13*	9	46.9 ³	07 00 ⁸	15
10	49.8 ³	02 26 ²	64	10	49.5 ⁵	09 43 ⁷	17
	20			11	50.4 ²	03 35 ⁵	11
1	15.2 ³	08 49 ⁸	15*	12	51.7 ³	00 54 ⁷	13
2	15.7 ³	01 54 ⁷	14	13	52.3 ³	02 43 ⁷	16
3	16.9 ³	04 00 ⁷	13	14	55.3 ⁴	08 08 ⁶	16
4	25.6 ⁶	06 22 ⁸	16	15	57.2 ⁴	09 43 ⁸	15
5	35.7 ⁵	04 13 ¹⁰	11		23		
6	37.3 ⁵	05 20 ⁸	14	1	05.1 ⁴	03 23 ⁶	9.4
7	39.1 ³	00 48 ⁷	12	2	08.2 ³	07 28 ⁶	22 ⁽²⁹⁾
8	42.4 ⁵	03 29 ⁷	11	3	09.6 ⁴	09 16 ⁸	51 (29) ⁽³⁰⁾
9	45.9 ³	01 54 ⁷	19	4	10.5 ³	04 50 ⁶	18
10	45.9 ³	06 57 ⁷	22†	5	14.0 ¹	03 53 ³	57
11	47.7 ³	04 00 ⁷	10	6	19.0 ⁴	09 16 ⁷	7.8*
12	55.4 ⁴	00 42 ⁴	104 (23)	7	24.9 ⁴	06 49 ⁸	15
13	55.7 ⁴	05 43 ⁸	19	8	25.0 ⁴	03 54 ⁶	17
	21			9	31.3 ⁴	01 03 ⁶	8.6
1	07.6 ⁴	06 19 ¹⁰	12* ⁽²⁷⁾	10	35.5 ⁷	05 39 ⁷	13
2	12.7 ³	04 06 ⁵	12	11	39.3 ³	04 38 ⁵	13†
3	21.9 ³	02 45 ⁶	17	12	57.1 ⁵	09 48 ⁵	12*

(25) Perhaps slightly extended. (26) Perhaps a galactic irregularity.

(27) A doubtful source. (28) May be several sources.

(29) (Pegasus I cluster). (30) (Pegasus II cluster).

TABLE 2
SOURCES BETWEEN DECLINATIONS 0° AND -10°

Sources observed on one record only are indicated by an asterisk. Sources which may be "extended", that is, resolvable, are indicated by a dagger. A colon has been placed beside uncertain flux densities

Ref. No.	Position (1950)		Flux Density (10^{-26} $\text{W m}^{-2} (\text{c/s})^{-1}$)	Ref. No.	Position (1950)		Flux Density (10^{-26} $\text{W m}^{-2} (\text{c/s})^{-1}$)
	R.A. h m	Dec. S. °			R.A. h m	Dec. S. °	
	00				02		
1	03.3 ³	00 56 ⁴	35	1	02.6 ⁴	05 33 ⁶	8.5
2	06.0 ⁴	06 19 ³	15	2	08.3 ⁴	03 38 ⁶	12 ⁽⁴⁾
3	17.4 ³	05 01 ⁶	9.5	3	10.7 ³	08 11 ⁶	8.5
4	17.7 ³	02 51 ⁴	23	4	10.8 ⁴	04 54 ⁶	8.7
5	18.8 ³	01 42 ⁶	9.8	5	12.4 ³	02 46 ⁶	7.5
6	21.5 ⁴	08 14 ⁶	24†	6	14.2 ³	00 54 ⁶	12
7	32.3 ²	08 27 ⁴	13 ⁽¹⁾	7	18.6 ²	02 11 ³	74 ⁽⁵⁾
8	32.6 ³	07 32 ⁶	9.0	8	18.6 ⁴	03 45 ⁶	7.5
9	36.4 ²	02 50 ⁵	120 (67) ⁽²⁾	9	29.4 ³	04 55 ⁵	12
10	39.0 ⁴	06 23 ⁵	10	10	29.8 ⁴	00 18 ⁶	14
11	39.2 ¹	09 43 ³	56	11	29.8 ³	06 57 ⁵	15
12	42.9 ³	00 05 ⁵	12	12	30.8 ⁴	02 40 ⁶	11
13	46.0 ³	07 01 ⁵	12	13	39.4 ⁴	02 30 ⁶	15
14	46.7 ³	02 48 ⁶	18	14	40.0 ¹	00 09 ³	35 ⁽⁶⁾
15	51.7 ³	03 42 ⁴	23	15	42.8 ³	05 21 ⁵	25†
16	52.4 ⁴	05 06 ⁶	8.5	16	43.6 ³	09 50 ⁷	8.7
17	54.5 ¹	01 39 ²	90 (72)	17	46.3 ⁴	07 46 ⁶	9.0
	01			18	54.1 ⁴	03 30 ⁶	11
1	06.5 ³	00 57 ⁶	13	19	56.8 ³	05 06 ⁶	8.8
2	10.5 ³	05 07 ⁶	15	20	57.8 ³	07 30 ⁵	11
3	19.6 ³	00 13 ⁵	19		03		
4	21.1 ⁴	03 50 ⁶	18	1	12.6 ³	03 37 ⁴	20
5	23.5 ¹	01 35 ²	88	2	29.8 ³	07 40 ⁵	12
6	28.8 ³	07 03 ⁶	19 ⁽³⁾	3	31.7 ²	01 25 ⁴	64
7	30.8 ³	00 26 ⁵	10	4	39.0 ³	04 55 ⁵	8.8 ⁽⁷⁾
8	35.1 ³	09 25 ⁴	18	5	46.0 ⁴	04 20 ⁷	16
9	35.4 ³	02 06 ⁵	13	6	49.6 ³	07 25 ⁶	25
10	43.7 ²	02 27 ⁵	12	7	56.5 ⁵	03 50 ⁶	11
11	45.5 ⁴	00 02 ⁶	12	8	59.2 ³	02 10 ⁶	16 ⁽⁸⁾
12	47.6 ⁴	09 09 ⁶	9.4				
13	49.9 ²	03 52 ⁴	20				
14	51.6 ⁴	07 26 ⁶	9.0				
15	52.2 ³	05 17 ⁷	6.8				
16	55.1 ³	00 39 ⁶	8.0				
17	57.0 ⁴	02 31 ⁶	8.5				

(1) (NGC 157). (2) May be two sources.

(3) (NGC 584). (4) Perhaps background irregularity.

(5) IAU 02S0A. (6) NGC 1068.

(7) (NGC 1417).

(8) Interpretation difficult, complex response.

TABLE 2 (Continued)

Ref. No.	Position (1950)		Flux Density (10^{-26} $\text{W m}^{-2} (\text{c/s})^{-1}$)	Ref. No.	Position (1950)		Flux Density (10^{-26} $\text{W m}^{-2} (\text{c/s})^{-1}$)
	R.A. h m	Dec. S. °			R.A. h m	Dec. S. °	
	04				05		
1	00.1 ³	09 56 ⁶	10	14	40.1 ⁴	05 16 ⁶	9.5 ⁽¹⁵⁾
2	00.9 ³	08 54 ⁶	19	15	45.6 ⁴	04 42 ⁸	6.1
3	05.5 ³	05 37 ⁵	12	16	46.6 ⁴	06 41 ⁶	9.0
4	06.0 ³	06 46 ⁶	17	17	48.0 ⁴	08 08 ⁶	15
5	09.5 ³	01 50 ⁵	15	18	52.0 ³	02 00 ⁶	29 ⁽¹⁶⁾
6	09.6 ¹	01 02 ⁴	35	19	53.1 ⁵	01 00 ⁶	19 ⁽¹⁷⁾
7	15.5 ³	05 35 ⁷	36 (18) ⁽⁹⁾	20	54.8 ³	03 27 ⁶	18 ⁽¹⁸⁾
8	15.7 ²	03 21 ⁴	28	21	56.8 ²	08 03 ⁵	14
9	20.3 ⁴	09 28 ⁶	8.5				
10	26.4 ⁴	01 15 ⁶	9.7				
11	28.2 ³	09 58 ⁶	7.3				
12	30.9 ²	08 44 ⁶	11	1	04.5 ³	04 02 ⁵	9.0
13	32.9 ³	05 30 ⁴	10	2	06.1 ²	07 21 ⁴	23†
14	39.0 ³	09 52 ⁵	17	3	12.0 ³	03 53 ⁵	15
15	39.6 ³	00 49 ⁶	12	4	25.0 ⁴	05 56 ³	120
16	46.8 ³	09 55 ⁵	16	5	27.7 ⁴	02 25 ⁶	8.7
17	47.3 ⁴	04 33 ⁶	46 (23) ⁽⁹⁾	6	38.9 ⁵	06 40 ⁸	9.5
18	49.3 ³	06 38 ⁴	9.6	7	39.0 ⁵	08 01 ⁶	50 (25)
19	49.6 ⁴	02 31 ⁵	13	8	45.0 ⁴	02 06 ⁶	33†
20	52.4 ⁴	00 24 ⁶	20 (12)	9	45.3 ⁴	08 10 ⁶	17
21	58.7 ⁴	03 39 ⁶	18	10	45.6 ⁴	09 16 ⁶	11
22	59.6 ³	05 48 ⁶	8.5	11	47.2 ³	05 37 ⁵	25
				12	56.7 ²	02 12 ⁵	24
	05				07		
1	00.0 ⁴	08 37 ⁶	9.1	1	07.0 ³	00 38 ⁷	11†
2	10.0 ³	07 36 ⁶	16	2	10.4 ³	09 06 ⁵	21
3	12.4 ³	02 19 ⁵	17†	3	12.7 ²	02 41 ⁴	25
4	13.0 ²	01 15 ⁶	18	4	22.3 ³	09 49 ⁴	36
5	13.3 ³	09 41 ⁶	8.8	5	23.1 ³	06 10 ⁶	94 (47)
6	18.3 ³	06 15 ⁵	17 ⁽¹⁰⁾	6	24.4 ²	02 00 ⁴	29
7	22.2 ⁴	02 46 ⁶	16	7	31.4 ⁴	05 31 ⁶	8.6
8	22.3 ⁴	07 22 ⁶	15 ⁽¹¹⁾	8	36.2 ³	02 03 ⁵	19
9	23.6 ³	09 36 ⁶	12	9	38.8 ³	01 01 ⁶	15
10	27.9 ³	00 03 ⁵	15	10	44.2 ²	08 05 ⁶	17
11	32.5 ²	05 24 ³	83 (69) ⁽¹²⁾	11	48.6 ³	06 52 ⁶	11
12	38.0 ⁵	02 20 ¹⁰	88 (24) ⁽¹³⁾	12	58.9 ⁴	02 06 ⁶	7.3
13	39.1 ⁴	01 25 ⁶	23 ⁽¹⁴⁾	13	59.7 ³	09 40 ⁶	17

⁽⁹⁾ May be two sources.^{(10), (11)} Perhaps one extended source.⁽¹²⁾ M 42.⁽¹³⁾ IC 434 etc.⁽¹⁴⁾ Possibly connected with source 05—012.⁽¹⁵⁾ (M 42—eastward extension).^{(16), (17), (18)} Perhaps part of Barnard's ring.

TABLE 2 (Continued)

Ref. No.	Position (1950)		Flux Density (10^{-26} $\text{W m}^{-2} (\text{c/s})^{-1}$)	Ref. No.	Position (1950)		Flux Density (10^{-26} $\text{W m}^{-2} (\text{c/s})^{-1}$)
	R.A. h m	Dec. S. °			R.A. h m	Dec. S. °	
	08				10		
1	01.0 ²	04 13 ⁶	14	7	23.1 ³	08 10 ⁶	11
2	03.1 ³	00 30 ⁶	15	8	24.1 ³	02 19 ⁶	17
3	03.9 ⁴	07 54 ⁷	9.7	9	24.3 ³	04 47 ⁷	5.3
4	09.3 ²	05 40 ⁵	22	10	25.4 ³	07 20 ⁶	10
5	13.4 ²	02 53 ⁵	35 ⁽¹⁹⁾	11	27.3 ²	05 57 ⁶	17†
6	21.5 ²	09 32 ⁶	20†	12	30.1 ⁴	09 10 ⁷	6.5†
7	22.7 ⁴	04 38 ⁷	8.8	13	33.4 ³	02 29 ⁶	16
8	27.2 ³	03 15 ⁶	27 (19)	14	33.7 ⁴	06 17 ⁷	6.5
9	32.0 ³	05 10 ⁶	13	15	36.0 ⁴	00 53 ⁵	8.2
10	32.3 ³	07 25 ⁶	13	16	41.9 ⁵	08 12 ⁷	17†
11	34.3 ³	01 04 ⁷	13	17	44.7 ³	01 06 ⁶	14
12	40.3 ⁴	09 15 ⁷	7	18	46.3 ²	02 33 ⁵	20
13	53.5 ³	06 07 ⁶	12	19	48.5 ⁴	09 19 ⁷	8.5*
14	57.6 ⁴	02 05 ⁷	7.8	20	59.6 ⁴	09 39 ⁶	6.5*
15	59.9 ³	05 07 ⁶	18	21	59.8 ²	00 52 ⁵	23
	09				11		
1	01.3 ³	06 46 ⁶	13	1	00.3 ³	06 18 ⁶	15
2	06.5 ³	09 38 ⁸	17 ⁽²⁰⁾	2	03.0 ⁴	08 22 ⁷	8.5
3	06.9 ⁴	03 15 ⁷	7.6	3	05.4 ⁴	03 55 ⁷	5.0
4	07.0 ³	01 22 ⁷	7.3	4	09.0 ³	06 10 ⁶	12
5	21.6 ⁴	04 22 ⁷	9.5	5	11.8 ³	01 54 ⁶	18
6	34.9 ³	04 00 ⁶	15	6	13.3 ³	07 10 ⁶	16
7	38.8 ³	01 17 ⁶	12	7	16.1 ³	08 43 ⁶	15
8	41.4 ⁴	07 14 ⁷	8.4	8	16.9 ²	02 46 ⁵	31
9	42.0 ⁵	09 39 ⁸	100 (28)* ⁽²¹⁾	9	25.4 ⁴	06 52 ⁷	14†
10	48.7 ³	04 57 ⁶	9.3	10	28.1 ³	03 15 ⁶	10
11	48.9 ³	08 31 ⁶	12	11	31.4 ⁴	07 43 ⁷	6.0
	10			12	34.2 ⁴	00 30 ⁷	8.2
1	05.3 ²	09 45 ⁵	17	13	39.3 ⁴	01 28 ⁷	6.3
2	07.3 ³	03 44 ⁵	10	14	41.6 ³	03 45 ⁵	8.4*
3	08.1 ³	07 25 ⁶	17	15	42.7 ⁵	06 06 ⁷	6.0
4	11.8 ⁴	09 30 ⁶	9.4	16	42.9 ²	00 12 ⁵	24†
5	16.9 ³	02 34 ⁶	16	17	46.4 ³	06 59 ⁶	16†
6	17.7 ⁴	03 00 ⁷	7.3	18	56.2 ³	00 30 ⁶	16

⁽¹⁹⁾ IAU 08S0A.⁽²⁰⁾ Perhaps together with 09—13 makes an extended source.⁽²¹⁾ Complex distribution, may be several sources.

TABLE 2 (Continued)

Ref. No.	Position (1950)		Flux Density (10^{-26} $\text{W m}^{-2} (\text{c/s})^{-1}$)	Ref. No.	Position (1950)		Flux Density (10^{-26} $\text{W m}^{-2} (\text{c/s})^{-1}$)
	R.A. h m	Dec. S. ° ' "			R.A. h m	Dec. S. ° ' "	
	12				14		
1	01.8 ³	04 36 ⁶	11.8	1	04.2 ²	02 09 ⁸	12†
2	03.7 ³	07 37 ⁷	18	2	05.5 ²	06 19 ⁵	18
3	04.4 ²	07 27 ⁵	9.9	3	06.1 ²	09 49 ⁸	10
4	05.0 ³	08 42 ⁶	11	4	06.5 ³	08 59 ⁸	27 (17)
5	08.6 ⁴	09 38 ⁷	11	5	09.0 ²	02 58 ⁸	7
6	11.2 ³	04 36 ¹⁰	9.9	6	09.6 ²	06 52 ⁷	14
7	11.9 ³	00 36 ⁸	15	7	14.7 ³	03 50 ⁷	24.4†
8	15.9 ³	04 47 ⁷	15.7	8	19.7 ³	05 20 ⁶	5.0
9	15.9 ²	09 54 ⁷	23	9	20.1 ³	09 09 ⁷	16
10	16.1 ²	07 03 ⁵	9.2	10	23.4 ⁶	08 00 ⁷	7.5
11	35.9 ³	00 31 ⁶	8.0 ⁽²²⁾	11	26.6 ⁴	01 18 ⁵	25
12	37.1 ⁵	07 19 ⁶	52 (24)	12	29.1 ³	03 38 ⁵	16
13	37.7 ³	04 24 ⁶	25	13	34.7 ²	08 21 ⁶	14
14	39.2 ²	08 33 ⁷	20	14	37.4 ³	06 56 ⁵	22†
15	40.5 ⁴	06 07 ⁸	12	15	42.4 ⁴	08 44 ⁷	20†
16	43.1 ⁴	03 06 ⁶	10	16	43.0 ³	03 45 ⁷	7.6
17	44.8 ³	05 21 ⁸	14	17	52.7 ²	04 10 ⁶	22
18	45.5 ⁴	06 12 ⁸	17	18	53.5 ²	05 44 ⁵	16
19	48.0 ²	01 36 ⁶	14	19	55.5 ³	00 54 ⁷	19 ⁽²⁴⁾
20	53.7 ¹	05 38 ⁵	37				
21	57.3 ²	00 24 ⁶	7.5				
	13				15		
1	04.2 ³	05 42 ⁸	8.5	1	02.6 ³	00 18 ¹³	18
2	06.0 ⁴	09 49 ⁷	19	2	04.3 ²	06 41 ⁶	17
3	07.3 ²	00 29 ⁵	25	3	08.3 ³	00 42 ⁶	19
4	09.6 ³	02 29 ⁷	11	4	09.0 ²	09 17 ⁵	18
5	12.8 ⁴	08 05 ⁸	45 (23)	5	09.0 ⁴	05 26 ⁶	8
6	13.0 ³	06 17 ⁶	7.0	6	09.2 ³	08 15 ⁷	7.5
7	13.1 ³	01 25 ⁸	16	7	20.4 ³	05 12 ⁷	14†
8	16.8 ⁴	00 30 ⁷	13	8	21.8 ³	06 52 ⁷	12†
9	28.4 ⁴	06 07 ⁶	13	9	21.9 ⁴	03 13 ⁶	7.0
10	33.8 ³	07 54 ⁷	8.7	10	22.1 ³	07 28 ⁷	18
11	35.7 ²	06 21 ⁶	35	11	22.9 ³	08 17 ⁶	12
12	41.8 ²	03 04 ⁷	16	12	38.1 ⁵	01 54 ⁶	37 (21)
13	43.0 ³	07 48 ⁷	53 (23)	13	39.0 ³	04 59 ⁷	12
14	48.1 ³	09 55 ⁷	8.0	14	42.5 ³	03 41 ⁹	23
15	48.3 ³	05 36 ⁶	12	15	45.3 ⁵	07 20 ⁹	14.8*
16	50.0 ³	06 07 ⁷	7.8 ⁽²³⁾	16	46.0 ²	07 55 ⁸	12
17	53.4 ²	08 10 ¹⁰	8.7	17	51.8 ⁵	02 52 ⁵	9.7
18	56.7 ⁴	09 57 ⁸	8.5	18	52.4 ²	06 57 ⁸	19
				19	53.4 ⁴	09 05 ⁷	10
				20	57.3 ³	04 38 ⁷	7.0*

⁽²²⁾ (NGC 4592).⁽²³⁾ (NGC 5324).⁽²⁴⁾ (NGC 5792).

TABLE 2 (Continued)

Ref. No.	Position (1950)		Flux Density (10^{-26} $\text{W m}^{-2} (\text{c/s})^{-1}$)	Ref. No.	Position (1950)		Flux Density (10^{-26} $\text{W m}^{-2} (\text{c/s})^{-1}$)
	R.A. h m	Dec. S. ° ' "			R.A. h m	Dec. S. ° ' "	
	16				18		
1	02.7 ³	09 15 ⁸	20	8	31.6 ²	08 42 ¹⁰	160
2	05.8 ³	06 36 ⁸	9.5	9	38.5 ²	05 10 ⁸	20:
3	12.3 ²	02 30 ⁴	9.1	10	41.7 ⁴	03 51 ⁵	180 (100)
4	12.4 ²	00 35 ⁵	15	11	41.8 ⁵	01 48 ¹⁰	25:
5	14.0 ³	05 44 ⁷	9.5	12	46.3 ⁴	00 53 ⁸	20:
6	16.0 ³	08 42 ⁷	8.0	13	50.3 ⁴	07 48 ⁸	17
7	26.4 ⁴	06 20 ⁶	9.0	14	53.0 ⁴	02 42 ⁷	150 (40) ⁽²⁸⁾
8	26.6 ⁴	03 24 ⁶	17	15	57.9 ³	04 13 ⁵	34
9	34.1 ³	03 33 ⁸	12				
10	36.9 ⁵	00 30 ⁶	26		19		
11	42.7 ²	07 14 ⁷	21	1	04.2 ⁴	03 06 ⁷	53 (34)
12	49.3 ⁴	00 18 ⁶	80 (50)	2	08.9 ³	06 41 ⁷	16
13	52.0 ⁴	05 09 ⁶	11	3	11.3 ³	09 41 ⁷	15†
14	52.6 ²	02 17 ⁹	60 (26)	4	14.1 ⁵	02 17 ⁷	28
15	54.6 ³	09 08 ⁶	11	5	18.5 ³	05 33 ⁸	9.8* ⁽²⁹⁾
16	56.0 ³	01 11 ⁷	15:	6	20.0 ⁶	03 38 ⁷	16
				7	26.2 ⁵	02 05 ⁵	23
				8	28.1 ³	06 41 ⁶	12
				9	32.6 ³	09 46 ⁸	23
				10	39.8 ³	04 36 ⁸	12
				11	40.8 ⁵	07 29 ⁶	33 (24)
				12	42.9 ⁴	04 55 ⁸	20
				13	43.5 ²	02 45 ⁶	22
				14	44.8 ⁴	00 13 ⁷	22
				15	45.8 ³	08 54 ⁸	9.5
				16	53.3 ³	05 22 ⁸	10
					20		
				1	06.4 ³	04 25 ⁷	19
				2	09.7 ³	09 00 ⁷	9.5
				3	18.7 ⁴	09 38 ⁷	9.0
				4	23.2 ⁴	01 18 ⁷	14
				5	27.0 ⁴	00 47 ⁷	8.7
				6	28.6 ²	08 09 ⁷	14
				7	33.5 ²	09 27 ⁸	14
				8	37.5 ³	02 54 ⁷	9.7
				9	44.1 ³	02 17 ⁷	18
				10	45.0 ³	07 59 ⁸	9.0
				11	45.1 ³	03 15 ⁷	27 (19)
				12	53.2 ⁴	06 52 ⁸	15
				13	58.8 ⁴	08 49 ⁷	17†
	17						
1	05.8 ⁴	01 36 ⁶	17				
2	06.2 ³	04 41 ⁵	12† ⁽²⁵⁾				
3	09.7 ⁴	00 26 ⁷	15: ⁽²⁶⁾				
4	12.5 ⁴	03 16 ⁷	21				
5	16.9 ³	04 25 ⁸	31				
6	18.1 ¹	00 55 ³	475 ⁽²⁷⁾				
7	22.1 ³	03 50 ⁹	16				
8	24.6 ⁴	08 21 ⁷	15†				
9	30.9 ²	05 10 ⁷	16*				
10	33.7 ³	06 52 ⁸	19				
11	37.7 ⁴	01 18 ⁸	39				
12	48.1 ³	02 06 ⁸	53				
13	53.8 ³	08 10 ⁸	21*				
14	54.4 ³	05 34 ⁶	55*				
15	55.7 ²	01 24 ⁶	50†				
	18						
1	02.4 ³	05 19 ⁷	25				
2	05.2 ³	00 59 ⁸	62†				
3	12.4 ³	05 59 ⁷	82 (48)				
4	14.9 ³	07 03 ⁸	30: ⁽²⁸⁾				
5	17.6 ⁴	09 32 ⁵	50 (25) ⁽²⁸⁾				
6	20.6 ²	01 34 ⁶	76 (55)				
7	25.3 ³	04 38 ⁸	40:†				

⁽²⁵⁾ May be two sources. ⁽²⁶⁾ Possibly side lobe of source 17-06.⁽²⁷⁾ Perhaps superimposed on extended source.⁽²⁸⁾ Perhaps a galactic irregularity. ⁽²⁹⁾ A doubtful source.

TABLE 2 (Continued)

Ref. No.	Position (1950)		Flux Density (10^{-26} $\text{W m}^{-2} (\text{c/s})^{-1}$)	Ref. No.	Position (1950)		Flux Density (10^{-26} $\text{W m}^{-2} (\text{c/s})^{-1}$)
	R.A. h m	Dec. S. ° ,			R.A. h m	Dec. S. ° ,	
	21				22		
1	00.2 ³	09 45 ⁶	12	12	29.2 ⁴	08 33 ⁶	15†
2	00.9 ³	04 02 ⁷	19†	13	33.3 ³	07 03 ⁸	8.0
3	02.1 ³	00 30 ⁷	11	14	36.9 ³	04 13 ⁶	16
4	05.4 ³	07 06 ⁵	17	15	43.5 ⁴	02 10 ⁷	14 ⁽³²⁾
5	10.5 ³	09 50 ⁷	11	16	45.0 ³	02 52 ⁷	20 ⁽³³⁾
6	13.7 ³	02 47 ¹⁰	28 (15)	17	49.2 ⁴	03 25 ⁸	9:
7	24.9 ³	05 35 ⁸	19 ⁽³⁰⁾	18	53.1 ⁵	06 37 ⁸	12
8	25.0 ⁴	06 36 ⁶	10 ⁽³¹⁾	19	53.9 ⁵	00 18 ⁶	32 (19)
9	25.7 ⁵	00 59 ⁷	15*	20	54.9 ²	01 16 ⁶	6.6
10	28.1 ⁵	09 15 ⁷	16†	21	55.3 ³	08 32 ⁸	13
11	31.9 ³	02 28 ⁶	9.0				
12	31.6 ³	01 16 ⁶	6		23		
13	38.0 ⁴	07 02 ⁸	12	1	01.7 ⁴	02 17 ⁸	7
14	40.6 ⁴	09 14 ⁷	7.0	2	02.6 ³	05 27 ⁶	10
15	41.7 ³	04 02 ⁸	12	3	02.8 ⁴	01 00 ⁸	9.5
16	43.8 ³	08 10 ⁷	13	4	03.6 ³	03 43 ⁶	14
17	50.7 ³	03 40 ⁷	6.0	5	05.5 ³	07 59 ⁷	6.7
18	54.2 ³	01 29 ⁸	15.6	6	07.5 ⁴	09 22 ⁸	9.0
19	56.3 ³	05 55 ⁸	11	7	12.6 ³	05 57 ⁶	6.7
20	57.7 ³	03 55 ⁸	13	8	15.6 ³	02 29 ⁷	9.8
				9	19.5 ³	09 16 ⁵	6.0
				10	24.3 ³	05 15 ⁶	35 (19)
	22			11	25.1 ³	02 22 ⁶	19
1	02.2 ³	08 43 ⁸	11	12	25.2 ⁴	08 10 ⁸	9.0
2	04.6 ³	09 16 ⁵	10	13	32.7 ³	04 59 ⁵	9.7
3	05.4 ³	05 30 ⁶	7	14	33.4 ⁵	00 19 ⁵	9.5 ⁽³⁴⁾
4	05.7 ³	03 27 ⁸	13	15	38.0 ³	00 08 ⁶	11
5	10.8 ³	09 29 ⁵	17	16	42.5 ³	05 22 ⁷	7.6
6	16.3 ³	03 46 ⁶	33 (18)	17	46.1 ³	03 36 ⁸	8.6
7	16.9 ³	00 42 ⁹	13	18	48.7 ³	04 21 ⁸	13
8	19.4 ⁴	08 43 ¹⁰	7.1	19	49.7 ³	08 10 ⁷	10
9	21.5 ¹	02 18 ³	60	20	49.9 ³	01 23 ⁷	18
10	23.1 ³	05 13 ⁵	30	21	51.3 ³	05 30 ⁷	9
11	24.5 ³	03 39 ⁸	9.6				

⁽³⁰⁾, ⁽³¹⁾ Perhaps one extended source.⁽³²⁾, ⁽³³⁾ Perhaps one extended source.⁽³⁴⁾ (NGC 7716).

TABLE 3
 SOURCES BETWEEN DECLINATIONS -10° AND -20°

Sources observed on one record only are indicated by an asterisk. Sources which may be "extended", that is, resolvable, are indicated by a dagger. A colon has been placed beside uncertain flux densities

Ref. No.	Position (1950)		Flux Density (10^{-26} $\text{W m}^{-2} (\text{c/s})^{-1}$)	Ref. No.	Position (1950)		Flux Density (10^{-26} $\text{W m}^{-2} (\text{c/s})^{-1}$)
	R.A. h m	Dec. S. ° '			R.A. h m	Dec. S. ° '	
	00				01		
1	00.0 ²	17 32 ⁵	28	13	36.9 ⁴	17 49 ⁶	10 ⁽²⁾
2	00.3 ³	15 28 ⁵	15	14	38.4 ⁴	18 25 ⁷	8.0 ⁽³⁾
3	00.6 ³	12 23 ⁶	12	15	40.4 ²	16 51 ⁴	28
4	05.6 ³	19 58 ⁶	17	16	45.6 ³	18 44 ⁷	16
5	09.2 ²	19 07 ⁵	13	17	47.6 ³	11 11 ⁶	10
6	12.4 ³	15 07 ⁸	34 (20)	18	47.9 ³	13 11 ⁷	8.1
7	15.9 ³	13 02 ⁵	52 (33) ⁽¹⁾	19	50.6 ⁴	14 54 ⁶	12
8	16.2 ²	10 46 ⁵	23	20	55.1 ³	10 45 ⁶	16
9	18.6 ³	19 11 ⁵	8.7	21	59.6 ³	11 47 ⁶	14
10	25.0 ⁴	16 48 ⁶	6.0				
11	25.3 ⁴	13 10 ⁷	13		02		
12	27.6 ³	11 50 ¹⁰	14	1	02.0 ⁴	19 43 ⁷	8.5
13	29.4 ³	15 33 ⁶	8.8	2	03.5 ³	18 16 ⁶	17
14	32.5 ³	16 50 ⁶	12	3	08.0 ³	11 18 ⁶	30 (19)
15	32.5 ³	18 14 ⁵	17	4	11.4 ³	16 02 ⁶	8.2
16	35.0 ⁵	12 35 ⁸	9.6	5	13.2 ¹	13 19 ³	42 ⁽⁴⁾
17	38.0 ⁴	13 13 ⁷	10	6	14.8 ⁴	17 58 ⁸	8.5
18	39.0 ³	15 44 ⁶	14†	7	22.9 ³	11 38 ⁸	13
19	43.5 ⁴	14 49 ⁶	9.0	8	26.5 ³	17 31 ⁶	19
20	45.8 ⁴	17 58 ⁷	8.9	9	30.8 ³	10 12 ⁵	17
21	48.6 ³	12 28 ⁵	18	10	35.4 ¹	19 42 ³	44
22	50.1 ³	19 53 ⁷	11	11	36.0 ⁴	14 45 ⁷	14
23	52.3 ⁴	16 19 ⁶	12	12	36.3 ⁴	18 20 ⁶	9.5
24	56.9 ³	13 40 ⁶	13	13	45.8 ³	16 47 ⁶	6.2
25	57.2 ³	15 22 ⁶	17	14	46.2 ⁴	13 29 ⁸	15
26	57.6 ³	17 24 ⁵	29	15	47.5 ³	18 10 ⁵	9.3
27	58.9 ³	14 30 ⁶	9.8	16	56.2 ³	16 52 ⁶	12
	01				03		
1	01.6 ²	12 27 ⁵	18	1	03.5 ³	12 21 ⁵	18
2	05.9 ¹	16 15 ²	53	2	05.4 ⁴	16 44 ⁶	17
3	07.2 ³	18 51 ⁶	9.0	3	07.5 ³	13 33 ⁷	16
4	08.2 ⁴	14 33 ⁶	16	4	15.1 ³	14 48 ⁶	9.5
5	11.7 ⁴	10 07 ⁶	7.8	5	27.9 ³	16 51 ⁵	16
6	14.5 ³	11 53 ⁶	11	6	31.1 ⁴	18 48 ⁶	12
7	16.8 ⁵	16 45 ¹⁰	13†	7	44.1 ²	11 13 ⁴	34
8	16.8 ³	19 00 ⁶	14	8	46.4 ⁴	13 08 ⁶	10
9	18.0 ²	15 34 ³	45	9	49.3 ²	14 38 ³	44
10	24.9 ³	12 10 ⁶	7.0	10	49.7 ²	10 08 ⁵	21
11	25.1 ²	14 13 ³	30	11	57.5 ³	16 20 ⁷	18
12	27.9 ³	15 38 ⁵	18†				

⁽¹⁾ Extended N.-S., may be two sources. ^{(2), (3)} May be one extended source.

⁽⁴⁾ IAU 0281A.

TABLE 3 (Continued)

Ref. No.	Position (1950)		Flux Density (10^{-26} W m $^{-2}$ (c/s) $^{-1}$)	Ref. No.	Position (1950)		Flux Density (10^{-26} W m $^{-2}$ (c/s) $^{-1}$)
	R.A.	Dec.			R.A.	Dec.	
	h m	° ' S.			h m	° ' S.	
	04				05		
1	05.0 ³	13 20 ⁸	14†	22	49.3 ³	10 32 ⁴	17
2	05.4 ¹	12 26 ³	31	23	51.0 ³	16 59 ⁶	8.5
3	08.9 ⁴	16 27 ⁶	10	24	51.7 ⁵	14 19 ⁷	8.7
4	11.4 ⁴	19 36 ⁷	9.4	25	51.9 ³	12 29 ⁶	9.5*
5	11.8 ⁴	11 26 ⁶	18	26	57.6 ³	16 50 ⁶	13
6	13.8 ³	15 22 ⁸	15*				
7	16.3 ³	18 13 ⁵	13				
8	23.0 ³	16 57 ⁶	14*	1	03.9 ³	10 45 ⁶	9.2
9	23.9 ³	12 07 ⁵	16	2	04.6 ⁴	17 49 ⁷	15
10	25.9 ³	11 38 ⁷	11	3	07.3 ⁴	14 40 ⁷	14†
11	27.2 ⁵	18 36 ⁸	9.0	4	14.8 ⁵	15 00 ⁷	19
12	32.0 ²	13 26 ⁵	38	5	17.8 ⁵	16 36 ¹⁰	63 (21)
13	32.9 ⁴	16 38 ⁶	15	6	20.3 ⁴	13 39 ⁸	9.5
14	36.9 ⁴	15 00 ⁷	7.3	7	25.8 ³	12 52 ⁸	16†
15	38.3 ³	12 10 ⁶	8.0	8	34.1 ⁴	15 46 ⁷	16
16	42.8 ⁵	18 52 ⁷	7.0	9	34.9 ³	13 44 ⁸	9.3
17	48.0 ³	17 34 ⁶	14	10	36.3 ³	16 50 ⁶	18
18	52.1 ⁴	19 07 ⁸	7.3	11	42.2 ⁴	10 19 ⁶	84 (27) ⁽⁶⁾
19	54.2 ³	11 51 ⁶	17	12	44.0 ²	15 33 ⁶	18
20	59.9 ²	12 16 ⁴	14	13	49.7 ⁵	12 43 ¹⁰	55 (11) ⁽⁷⁾
				14	53.2 ³	19 15 ⁷	7.6†
	05				07		
1	03.0 ³	10 13 ⁵	20	1	03.2 ²	11 02 ⁷	55 (25) ⁽⁸⁾
2	06.5 ³	14 29 ⁶	16	2	03.6 ⁴	19 13 ⁷	10
3	08.5 ²	18 42 ³	41	3	12.0 ³	14 30 ¹⁰	17†
4	13.0 ³	15 56 ⁸	11	4	13.8 ⁴	11 20 ⁵	25†
5	13.6 ²	13 41 ⁶	16	5	16.2 ⁴	17 07 ⁷	17†
6	15.5 ²	16 34 ⁵	16	6	21.4 ³	18 38 ⁵	19
7	21.2 ⁴	11 59 ⁶	11	7	23.8 ³	13 16 ⁷	13†
8	23.8 ³	18 24 ⁶	14	8	26.1 ²	14 51 ⁶	17
9	24.2 ³	13 36 ⁶	16	9	29.7 ⁴	18 17 ⁸	29 (17)
10	24.9 ³	16 31 ⁷	12	10	32.9 ³	15 59 ⁶	12
11	24.9 ³	17 35 ⁶	8.2	11	34.2 ³	19 38 ⁶	11
12	25.4 ³	10 45 ⁶	16	12	34.8 ³	15 00 ⁷	9.2
13	26.6 ⁴	14 48 ⁷	8.3	13	38.6 ³	13 58 ⁶	12
14	33.3 ³	12 01 ⁶	15	14	41.5 ⁴	17 43 ⁷	9.8
15	34.6 ⁴	18 31 ⁸	12	15	43.4 ⁵	16 32 ⁷	10
16	35.0 ⁴	17 18 ⁸	15	16	45.6 ⁴	10 01 ⁶	13
17	35.3 ³	13 16 ⁸	14 ⁽⁵⁾	17	45.5 ²	19 00 ⁴	52
18	37.1 ³	16 04 ⁸	9.7	18	46.2 ⁴	11 53 ⁷	20
19	42.0 ⁴	12 33 ⁸	8.0	19	51.3 ⁵	19 22 ⁸	17
20	43.7 ³	17 33 ⁷	17				
21	48.7 ³	15 48 ⁶	8.7				

(5) Rather doubtful because of a large side lobe of 05N2A at this declination.

(6), (7) Perhaps one extended object elongated parallel to the galactic circle.

(8) Extended source with apparently complex brightness distribution. May be associated with the emission nebulae IC 2177 and NGC 2327.

TABLE 3 (Continued)

Ref. No.	Position (1950)		Flux Density (10^{-26} $\text{W m}^{-2} (\text{c/s})^{-1}$)	Ref. No.	Position (1950)		Flux Density (10^{-26} $\text{W m}^{-2} (\text{c/s})^{-1}$)
	R.A.	Dec. S.			R.A.	Dec. S.	
	h m	° ' ,			h m	° ' ,	
	08				10		
1	00.3 ⁴	14 40 ⁷	33 (18)*	4	10.1 ³	18 15 ⁶	14
2	03.4 ³	17 11 ⁵	18*	5	10.3 ⁴	15 16 ⁷	9.2 ⁽¹⁷⁾
3	05.3 ³	12 37 ⁶	14*	6	18.9 ³	19 43 ⁶	7.5
4	07.0 ²	10 27 ³	40	7	19.9 ³	10 25 ⁷	6.5*
5	13.1 ⁴	11 49 ⁷	4.2	8	22.4 ³	10 43 ⁶	18
6	13.8 ³	15 57 ⁵	14*	9	23.0 ⁴	11 44 ⁷	8.5
7	17.6 ⁴	11 00 ⁷	8.9 ⁽⁹⁾	10	23.6 ⁴	18 10 ⁶	10*
8	27.3 ⁴	17 39 ⁶	14	11	28.0 ³	15 28 ⁶	18
9	33.1 ⁴	16 04 ⁷	8.8*	12	30.0 ⁴	13 36 ⁷	7.5
10	35.0 ³	11 27 ⁶	18	13	31.0 ⁴	17 04 ⁸	9.0
11	39.6 ³	17 49 ⁶	10 ⁽¹⁰⁾	14	32.4 ⁴	19 15 ⁷	11
12	43.8 ³	11 28 ⁶	12	15	33.5 ³	10 20 ⁵	9.4
13	44.6 ⁴	17 44 ⁷	9.4 ⁽¹¹⁾	16	34.7 ³	18 24 ⁶	14
14	45.6 ⁴	15 33 ⁷	6.6 ⁽⁹⁾	17	38.7 ⁴	11 53 ⁷	6.5
15	48.4 ⁴	10 15 ⁷	7.6	18	39.4 ⁴	14 00 ⁷	9.3
16	51.3 ²	14 18 ⁵	24	19	44.8 ⁴	17 08 ⁷	7.5†
17	53.2 ³	12 27 ⁶	13	20	46.6 ³	18 46 ⁶	24†
18	54.4 ⁵	15 38 ⁸	9.4*	21	48.7 ³	20 12 ⁶	13
19	55.6 ³	19 38 ⁷	17	22	54.6 ⁴	16 00 ⁷	9.2
	09				11		
1	00.0 ³	14 18 ⁶	12	1	00.6 ⁵	15 01 ¹⁰	56 (14)*
2	03.5 ³	12 32 ⁶	16	2	10.4 ³	11 50 ⁶	10
3	06.4 ⁴	10 22 ⁷	9.5 ⁽¹²⁾	3	11.2 ³	13 15 ⁶	17*
4	15.7 ¹	11 53 ²	690 ⁽¹³⁾	4	19.9 ³	12 00 ⁶	12*
5	30.0 ³	19 56 ⁶	11	5	30.4 ⁴	15 16 ⁶	9.4*(18)
6	31.4 ³	16 47 ⁶	13	6	30.9 ²	19 22 ⁴	32
7	38.2 ⁴	17 18 ⁶	15†	7	32.6 ³	17 25 ⁶	19
8	39.3 ³	16 09 ⁷	10*	8	36.5 ¹	13 41 ⁴	44 ⁽¹⁹⁾
9	39.7 ³	11 28 ⁵	50 (25)	9	39.8 ³	17 11 ⁶	7.3
10	42.7 ³	19 33 ⁶	12	10	40.0 ²	15 08 ⁵	25
11	43.5 ²	13 19 ⁵	25 (16)*	11	40.3 ³	11 29 ⁶	14
12	47.0 ⁴	18 15 ⁷	12	12	42.6 ³	15 43 ⁶	15
13	53.3 ⁴	12 50 ⁷	9.7 ⁽¹⁴⁾	13	47.1 ³	11 47 ⁶	17
14	54.0 ³	13 36 ⁶	14 ⁽¹⁵⁾	14	50.4 ⁴	10 10 ⁷	7.7
	10			15	52.0 ⁴	15 22 ⁶	6.6
1	03.8 ⁴	10 38 ⁷	7.3	16	53.1 ³	17 39 ⁶	9.5
2	07.6 ³	11 47 ⁷	32 (16)	17	56.6 ⁴	11 42 ⁷	7.3
3	08.3 ³	14 47 ⁶	17 ⁽¹⁶⁾	18	59.5 ²	18 41 ⁶	10 ⁽²⁰⁾
				19	59.9 ³	10 27 ⁵	16

(9) A doubtful source. (10), (11) Perhaps one extended source.

(12) Perhaps one extended source with 09—02. (13) IAU 09S1A.

(14), (15) Perhaps one extended source. (16), (17) Perhaps one extended source.

(18) A rather doubtful source. (19) IAU 11S1A.

(20) NGC 4038/39.

TABLE 3 (Continued)

Ref. No.	Position (1950)		Flux Density (10^{-26} $\text{W m}^{-2} (\text{c/s})^{-1}$)	Ref. No.	Position (1950)		Flux Density (10^{-26} $\text{W m}^{-2} (\text{c/s})^{-1}$)
	R.A. h m	Dec. S. °			R.A. h m	Dec. S. °	
	12				14		
1	01.8 ³	15 33 ³	14	1	01.3 ⁴	19 23 ⁷	14†
2	02.4 ⁵	17 39 ¹⁰	48 (16) ⁽²¹⁾	2	09.8 ⁴	18 41 ⁷	15
3	04.0 ³	12 53 ¹⁰	56 (20)	3	15.3 ⁴	17 15 ¹⁰	14†
4	09.1 ²	10 55 ⁵	10	4	16.0 ²	15 47 ⁸	34 (22)
5	09.3 ³	19 27 ⁵	11†	5	17.7 ³	19 14 ⁸	11
6	13.7 ³	14 39 ⁷	6.3*	6	20.4 ²	14 29 ⁸	26 (16)*
7	18.2 ²	16 30 ⁵	12	7	20.5 ⁵	13 14 ⁸	12
8	22.5 ³	19 32 ⁸	9.0*	8	20.9 ³	18 20 ¹⁰	9.0
9	23.4 ²	11 22 ⁸	16	9	23.6 ⁴	17 28 ⁸	11
10	28.4 ¹	16 59 ⁴	38 ⁽²²⁾	10	24.6 ²	11 44 ⁴	22
11	34.0 ³	14 13 ⁷	9.6	11	31.4 ⁵	19 13 ⁸	8.9*
12	35.2 ³	19 53 ⁸	24 (12)*	12	32.0 ⁴	12 22 ⁶	6.5
13	37.3 ⁵	15 38 ⁶	14†	13	32.8 ³	11 11 ⁸	8.5*
14	41.9 ²	19 36 ⁵	18*	14	37.2 ²	17 08 ⁸	11
15	43.3 ³	17 50 ⁵	7.0	15	41.7 ²	18 00 ⁸	11†
16	43.6 ²	11 06 ⁵	18	16	42.9 ²	19 23 ⁶	14
17	51.6 ²	18 20 ⁷	13	17	44.1 ³	11 36 ⁷	17*
18	52.3 ¹	12 19 ⁴	53 ⁽²³⁾	18	46.9 ⁴	15 53 ⁸	42 (20)
19	57.0 ⁴	17 16 ⁶	27 (14)	19	50.2 ³	12 58 ⁶	19
20	58.1 ³	11 17 ⁵	19	20	51.7 ³	18 30 ⁷	9.3
				21	53.4 ²	11 02 ⁵	41†
				22	59.2 ⁵	19 53 ³	16*(25)
	13				15		
1	00.0 ³	18 03 ⁸	18	1	00.3 ²	14 41 ⁷	13*
2	08.3 ³	19 53 ⁷	7.0*	2	02.7 ²	12 00 ¹⁰	9.3
3	12.0 ³	12 07 ⁷	8.7	3	03.3 ³	16 36 ⁸	10
4	12.8 ²	18 41 ⁵	22	4	04.5 ²	13 52 ⁷	13*
5	31.7 ⁶	14 18 ¹⁰	13	5	08.1 ³	18 05 ⁸	15
6	31.9 ³	10 00 ⁷	18	6	10.6 ⁵	19 23 ⁶	49 (30)
7	34.4 ³	10 57 ⁷	17	7	14.1 ³	13 58 ⁷	19*
8	34.7 ³	17 55 ⁶	11 ⁽²⁴⁾	8	16.6 ²	12 32 ⁶	13
9	41.4 ⁴	19 22 ⁸	14	9	23.5 ²	13 41 ⁴	16*
10	41.7 ⁴	12 21 ⁶	18*	10	27.1 ³	12 21 ⁶	8.2
11	45.4 ³	11 07 ⁷	15	11	31.5 ³	18 36 ⁸	13
12	46.8 ⁴	12 58 ⁷	14	12	37.8 ²	17 23 ⁷	16†
13	47.2 ⁴	16 30 ⁵	12	13	40.9 ⁵	16 02 ⁸	7.5*(25)
14	52.1 ³	19 23 ⁵	15	14	41.3 ³	13 36 ¹⁰	8.8*
15	53.9 ²	17 39 ⁶	18†	15	43.9 ²	12 23 ⁷	9.5
16	56.8 ³	16 17 ⁷	8.7	16	48.6 ²	19 51 ⁵	11
17	59.1 ²	11 35 ⁵	13	17	50.0 ³	16 57 ¹⁰	21 (14)
18	59.9 ²	14 50 ⁹	15†	18	53.3 ³	16 10 ⁷	10†

⁽²¹⁾ Possibly several sources.⁽²²⁾ Possibly a side lobe of IAU 12N1A, but appears genuine.⁽²³⁾ (NGC 4783), (NGC 4782). ⁽²⁴⁾ (NGC 5247). ⁽²⁵⁾ A doubtful source.

TABLE 3 (Continued)

Ref. No.	Position (1950)		Flux Density (10 ⁻²⁶ W m ⁻² (c/s) ⁻¹)	Ref. No.	Position (1950)		Flux Density (10 ⁻²⁶ W m ⁻² (c/s) ⁻¹)
	R.A. h m	Dec. S. ° '			R.A. h m	Dec. S. ° '	
	16				18		
1	03.2 ³	17 19 ⁶	16*	1	00.1 ⁵	17 49 ⁷	40:
2	04.1 ³	18 20 ¹⁰	7.6	2	04.7 ⁵	11 26 ⁷	29†
3	05.5 ³	16 18 ⁸	8.5	3	11.6 ²	17 12 ³	160:
4	07.7 ³	12 45 ⁷	15	4	12.0 ⁴	12 40 ¹⁰	20:
5	08.1 ⁴	10 44 ⁷	11	5	14.9 ³	10 57 ⁷	35:†
6	16.9 ⁴	10 05 ⁷	17	6	18.9 ⁵	18 38 ¹⁰	15:
7	17.6 ³	13 36 ⁶	12*	7	21.5 ³	13 50 ⁵	40:
8	21.1 ³	11 28 ⁴	20	8	21.8 ³	12 24 ⁴	150
9	22.0 ³	17 34 ⁷	15*	9	25.0 ³	11 17 ⁴	50
10	22.8 ⁵	19 23 ⁵	11	10	26.5 ³	17 54 ⁷	15:
11	30.4 ²	12 48 ⁶	15*	11	27.5 ³	12 46 ⁸	40:
12	32.6 ⁵	15 18 ⁹	14*	12	28.7 ³	14 36 ⁸	30:
13	34.9 ³	14 18 ⁷	16*	13	30.1 ²	10 01 ⁴	230
14	36.9 ⁵	12 53 ⁷	8.9*	14	42.1 ⁴	19 40 ⁸	56: (28)
15	38.0 ²	19 35 ⁶	23*	15	42.9 ⁴	13 37 ⁸	24†
16	38.1 ⁵	17 50 ¹⁰	19	16	48.9 ³	10 55 ⁷	23
17	40.4 ³	15 19 ⁵	30:*	17	51.1 ³	17 08 ⁷	15†
18	43.1 ⁴	18 20 ⁶	18				
19	45.4 ²	10 48 ⁸	37*				
20	48.1 ³	12 53 ⁷	14*				
21	55.5 ⁴	18 51 ⁸	17*				
22	55.7 ³	14 03 ⁵	22*				
	17				19		
1	05.2 ³	10 02 ⁶	15*	1	04.9 ⁵	19 01 ⁹	20
2	05.4 ³	17 13 ⁶	60 (35)*	2	05.8 ³	12 37 ⁵	17 ⁽³⁰⁾
3	10.5 ⁵	13 41 ⁸	32*†	3	11.3 ³	15 11 ⁶	17
4	15.0 ⁷	12 43 ⁷	16*†	4	14.7 ³	16 30 ⁸	12
5	15.9 ⁴	16 25 ⁷	15*	5	14.9 ²	11 58 ⁶	25
6	19.4 ⁵	18 45 ⁷	150 (50)* ⁽²⁶⁾	6	24.1 ⁵	14 18 ⁹	28
7	22.3 ⁴	10 49 ⁸	21*	7	27.1 ²	15 19 ⁶	23
8	37.1 ⁴	11 40 ⁶	16*	8	29.5 ²	19 44 ⁷	22
9	47.7 ⁴	13 04 ⁹	18*	9	31.7 ³	17 18 ⁸	12
10	48.7 ⁴	17 28 ⁸	30:*† ⁽²⁷⁾	10	32.2 ⁴	10 55 ⁷	75 (34)
11	51.1 ³	14 56 ⁹	19*	11	37.7 ³	15 36 ⁴	38:
12	51.3 ⁵	10 43 ⁸	16* ⁽²⁸⁾	12	39.7 ³	13 26 ⁷	13
13	53.9 ³	11 39 ⁶	12* ⁽²⁹⁾	13	48.9 ³	14 08 ⁸	15
14	55.4 ³	16 07 ⁷	24*	14	49.9 ³	18 10 ⁷	11
				15	50.6 ⁴	19 43 ⁷	18:†
				16	53.3 ⁴	12 30 ⁷	19
				17	54.1 ³	16 30 ⁶	9.2

⁽²⁶⁾ Perhaps two sources.⁽²⁷⁾ Measurements doubtful because of side lobe difficulties.^{(28), (29)} Perhaps one extended source.⁽³⁰⁾ A doubtful source.

TABLE 3 (Continued)

Ref. No.	Position (1950)		Flux Density (10^{-26} $\text{W m}^{-2} (\text{c/s})^{-1}$)	Ref. No.	Position (1950)		Flux Density (10^{-26} $\text{W m}^{-2} (\text{c/s})^{-1}$)
	R.A. h m	Dec. S. ° ' "			R.A. h m	Dec. S. ° ' "	
	20				21		
1	04.1 ³	19 32 ⁸	15	16	35.2 ³	18 54 ⁷	23
2	08.2 ³	16 14 ⁸	8.3	17	38.2 ³	16 35 ⁶	16
3	21.3 ³	17 38 ⁸	8.5	18	46.2 ³	17 07 ⁸	13†
4	21.9 ³	13 56 ⁹	6.7*	19	46.9 ⁴	13 36 ⁷	25 (13)
5	22.4 ⁴	19 43 ⁷	8.8†	20	48.7 ³	15 54 ⁷	8.8*
6	25.5 ²	15 41 ⁴	20	21	48.9 ³	19 53 ⁷	18†
7	33.2 ³	17 54 ⁶	15	22	53.7 ²	12 53 ⁷	8.8
8	36.5 ⁴	13 47 ⁷	13	23	54.2 ³	18 25 ⁶	25
9	40.9 ⁴	15 00 ⁸	9.0	24	58.2 ²	17 04 ⁵	14
10	43.0 ⁴	10 12 ¹⁰	8.0	25	58.5 ⁴	13 30 ⁶	12 ⁽³¹⁾
11	45.0 ³	18 20 ⁷	15				
12	48.5 ³	14 45 ⁷	13		22		
13	48.8 ³	16 15 ⁹	17	1	03.0 ²	18 40 ⁵	16:
14	50.2 ³	16 23 ⁷	13	2	03.4 ³	15 33 ¹⁰	6.7
15	50.3 ³	18 41 ⁷	9.0	3	07.6 ³	14 13 ⁵	10
16	53.5 ³	12 22 ⁷	8.5	4	08.5 ³	10 12 ⁶	9.5
17	56.8 ²	15 00 ⁶	13	5	08.5 ³	12 58 ⁷	14
18	58.2 ³	17 48 ⁶	24	6	10.3 ³	11 58 ⁶	16
19	59.7 ³	13 20 ⁷	14	7	12.0 ¹	17 11 ⁴	127 ⁽³²⁾
				8	21.4 ³	15 43 ⁶	10
				9	22.6 ³	14 08 ⁶	15
				10	23.0 ²	16 46 ⁶	15
				11	27.1 ³	18 51 ⁷	11†
				12	28.0 ³	10 23 ⁸	6.5
				13	34.9 ²	13 56 ⁶	10
				14	35.4 ²	12 03 ⁷	16
				15	35.8 ²	17 36 ⁶	17
				16	36.7 ³	19 33 ⁷	17
				17	39.9 ⁴	14 56 ⁷	6.0
				18	40.6 ⁴	16 36 ⁷	8:
				19	43.7 ²	19 02 ⁵	8.0
				20	56.0 ²	12 11 ⁶	8.6
				21	56.9 ³	15 12 ⁸	12
				22	57.4 ³	13 35 ⁸	6.7
				23	58.0 ⁵	10 28 ⁸	8.0
	21						
1	01.4 ²	10 44 ⁵	14				
2	03.4 ²	11 28 ⁶	12				
3	07.3 ³	13 25 ⁸	10				
4	15.3 ³	16 03 ⁷	9.3				
5	15.8 ⁴	14 08 ⁷	14				
6	17.0 ³	12 02 ⁶	7.1				
7	17.7 ²	15 16 ⁷	17				
8	19.1 ³	18 40 ⁸	9.7				
9	20.2 ⁴	16 49 ⁷	30:†				
10	24.5 ⁴	19 27 ⁸	8.2				
11	25.9 ³	12 11 ⁵	15				
12	26.0 ³	14 37 ⁷	8.4				
13	32.7 ³	13 09 ⁷	15				
14	33.3 ²	11 39 ⁶	28				
15	34.7 ²	14 39 ⁵	33				

⁽³¹⁾ (NGC 7171).⁽³²⁾ Perhaps slightly extended.

TABLE 3 (Continued)

Ref. No.	Position (1950)		Flux Density (10^{-26} $\text{W m}^{-2} (\text{c/s})^{-1}$)	Ref. No.	Position (1950)		Flux Density (10^{-26} $\text{W m}^{-2} (\text{c/s})^{-1}$)
	R.A.	Dec. S.			R.A.	Dec. S.	
	h m	° ' "			h m	° ' "	
	23				23		
1	04.8 ⁴	12 01 ⁷	8.6	14	26.7 ²	19 37 ⁵	19
2	06.5 ³	19 53 ⁷	11	15	27.3 ²	17 56 ⁶	11
3	07.7 ³	10 45 ⁷	7.6	16	27.6 ²	18 47 ⁶	13
4	09.6 ³	12 54 ⁶	11	17	29.2 ⁴	16 51 ⁶	10
5	13.9 ³	14 18 ⁷	9.6	18	30.0 ³	10 16 ⁷	10
6	14.1 ³	12 10 ⁶	8.6	19	34.9 ³	14 52 ⁶	16
7	15.9 ⁵	11 07 ⁷	6.9*	20	39.5 ⁴	12 51 ⁷	6.9
8	17.6 ³	16 30 ⁵	23	21	39.7 ³	16 46 ⁶	16
9	18.1 ³	19 32 ⁶	15	22	42.9 ⁴	15 22 ⁸	13
10	18.5 ³	13 36 ⁸	7.4	23	48.1 ⁴	16 25 ⁶	13
11	20.1 ⁴	15 33 ⁸	10	24	54.5 ³	13 20 ⁸	8.3
12	22.6 ²	12 29 ⁵	30	25	59.6 ²	17 26 ⁶	14
13	25.3 ³	15 02 ⁷	14				

have been made between different sets of measurements on the same sources and it appears that our estimates do represent the probable errors reasonably well; they may indeed be rather pessimistic, but the effects of systematic error or confusion due to finite resolutions would not show up in such a comparison. The Right Ascensions have been given only to the nearest 0.1 min, since this is the highest accuracy justified in the great majority of cases. As shown by Mills *et al.* (1958), there is probably a collimation error of -4° in Right Ascension, so for the most accurate positions, having a probable error of $\pm 0^{\text{m}}.1$, a correction of $+4^{\circ}$ has been applied before selecting the nearest 0.1 min. For some of the stronger sources the position accuracy is somewhat better than indicated in the catalogue; they will be discussed elsewhere.

Sources resolved by the aerial beam have been treated as before, both their peak flux density and their integrated flux density being given, the former in parentheses. Possible identifications with bright nebulae are also indicated and discussed in the next section.

Every effort has been made to produce a uniform treatment of the whole area, but this has not been possible for the weakest sources because of sensitivity variations. For instance, at the northern border of the catalogue zone, the sensitivity is only half that at the southern border. This has been compensated to some extent by taking more observations at the most northerly declinations, but simple inspection of the catalogue shows that there are many more faint sources listed in the southern zone. Similarly, close to the galactic circle at the crossing near the centre, the sensitivity is reduced by a magnitude or more, because of the very high brightness temperature of the galactic disk and the great complexity of the brightness distribution. However, it is considered that the

catalogue of Class II sources (i.e. those for which $|b| > 12\frac{1}{2}^\circ$) is complete down to a flux density of $2 \times 10^{-25} \text{ W m}^{-2} (\text{c/s})^{-1}$ over the whole area, and to a much lower level in regions of low sky temperature in southerly declinations.

III. IDENTIFICATIONS

A study of possible identifications of the radio sources with visible nebulae is in progress, using the Palomar Sky Atlas. This is a large undertaking which, to be really useful, requires the collection of information about any suspected nebula and, in addition, the measurement of the angular size of the associated radio source. As an interim measure we will therefore merely list some of the brighter sources which may possibly be identified with faint galaxies on the Palomar prints and, as in paper I, discuss in some detail possible identifications with objects listed in the Skalnate Pleso Catalogue (Becvar 1951) for which additional data are usually available.

Objects examined in the Skalnate Pleso Catalogue include emission nebulae, novae, planetary nebulae, globular clusters, and galaxies. The situation as regards emission nebulae is not different from that described in paper I and will not be enlarged upon. No additional identifications were obtained in the central galactic crossing now included in our catalogue, but this is not surprising in view of the high background brightness (e.g. Mills, Little, and Sheridan 1956). Of the novae, planetary nebulae, and globular clusters, the only coincidence in position is between the globular cluster NGC 7089 and the radio source 21-012. However, the position agreement is not at all close and, in view of the lack of radio emission from other globular clusters, it cannot be regarded as significant.

In the case of the galaxies the situation is different; two reasonably certain identifications have been made, and, on a statistical basis, it appears probable that there are several "radio galaxies" among a total of 20 coincidences noted in the catalogue. These coincidences are noted where source and galaxy have positions within 1^m in Right Ascension and $20'$ in declination, which correspond approximately to three times the mean probable errors in each coordinate. There are 315 galaxies listed in the area of the catalogue and it is easily shown that the number of coincidences expected is about 9 if galaxies and sources are completely uncorrelated. The number of actual coincidences, 20, is therefore significantly greater, but not enough to warrant further analysis. To reduce the chance coincidences we therefore restrict attention to those galaxies within $0^m.7$ in Right Ascension and $13'$ in declination, that is, within twice the mean probable error in each coordinate; there are 14 such coincidences and less than 5 expected by chance. The coincidences are listed in Table 4 together with our estimate of the quality of position agreement and the difference between radio and optical magnitudes, $m_{1.9} - m_p$, on the assumption that the source may be identified with the galaxy. The radio magnitude is defined by

$$m_\lambda = -53.4 - 2.5 \log S_\lambda,$$

and, for comparison with other data, it is converted to the scale of Brown and Hazard (1952) at 1.9 m by addition of the factor, 0.8 magnitude (Mills 1958, in press). The photographic magnitude used is the *total* magnitude as listed by de Vaucouleurs (1953).

The number of good and very good coincidences is the same as the difference between the actual and chance coincidences, suggesting that most may be real. The galaxies NGC 1068 and NGC 4038/4039 are almost certainly associated with the coincident radio sources, as the radio emission in each case is only about 2 magnitudes greater than the mean of the "normal" galaxies for which $m_{1.9} - m_p \simeq +1$ and in each case the extra emission is compatible with peculiarities in the galaxies; e.g. NGC 1068 is known to have strong and broad gaseous emission lines in the nucleus and the galaxies NGC 4038 and NGC 4039 appear to be undergoing a mild collision. No abnormalities are listed among the other galaxies, but this means little, for there seems no reason why an optically normal

TABLE 4
POSSIBLE IDENTIFICATIONS WITH BRIGHT GALAXIES

Radio Source	Galaxy		Agreement in Position	$m_{1.9} - m_p$
	NGC	Type		
00-07	157	Sc	Very poor	-1.3
01+03	470	Sbc	Fair	-3.5
01+03	474	E0	Good	-4.0
01+04	533	E3	Very good	-3.6
01-06	584	E3-4	Very good	-1.9
02-014	1068	Sb _p	Very good	-1.0
03-04	1417	S:	Good	-2.3
11-118	4038/39	Sc _p	Very good	-0.7
12+04	4234	I	Good	-3.6
12+05	4261	E2-3	Very good	-3.8
12-118	4782	SO:	Very poor	-4.7
12-118	4783	SO:	Poor	-4.7
14-019	5792	S	Very good	-2.5
21-125	7171	SBb	Very good	-2.3

galaxy should not emit substantially more than normal at radio frequencies; an established example is NGC 1316. In one case, NGC 7171, for which a Palomar print is available, there is an adjacent cluster of faint galaxies with several close to the radio position, one or more of which could well be the radio source; angular size measurements would clearly be useful here.

If the coincidences are taken at their face value, the suggestion is clear that a substantial proportion of galaxies may have a slightly abnormal radio emission. On the figures quoted, out of the 315 galaxies in the catalogue area, about 1 in 30 emits between 2 and 5 magnitudes more than a normal galaxy. While not very much weight can be given to this conclusion, it in no way contradicts the observed lack of correspondence between the majority of the sources and the brighter galaxies. It is necessary to increase the statistical reliability of the identifications by extending the catalogue to a larger area and, if possible, by increasing the sensitivity.

Our prints of the Palomar Sky Atlas are, at present, very incomplete, so that a systematic examination of the catalogue area is not yet possible. However,

it is worth noting a number of strong sources which correspond with the positions of galaxies on the Palomar prints down to a magnitude of about 18. These are the sources 02-15, 09+07, 12-110, 13-011, 16+02, 17-06, 21-125, and 23-112. Sources in an area between 00^h and 06^h near the celestial equator, which has been investigated by Minkowski, have not been included in this list. While positive identifications are not suggested for the above sources, they undoubtedly need further investigation.

We may also mention the IAU sources 09S1A (09-14) and 16N0A (16+010) for which Minkowski has suggested possible identifications. The former may possibly be identified with a faint double galaxy at position $09^h 15^m 42^s$, $-11^\circ 53'$ (1950); our most accurate radio position, after allowing for a probable collimation error of $1'$ in the aerial, is $09^h 15^m 44^s \pm 3^s$, $-11^\circ 52' \cdot 5 \pm 2'$ (1950). The angular size of the galaxy is about $\frac{1}{2}'$ compared with $1\frac{1}{2}'$ for the radio source (Carter 1955). Minkowski (unpublished data) notes that the $\lambda 3727$ line of [O II] in emission is detectable in the galaxy, but is not unusually strong. The corrected position of the radio source 16N0A is $16^h 48^m 46^s \pm 3^s$, $+5^\circ 04' \pm 2'$ (1950) compared with the position of a faint galaxy at position $16^h 48^m 49^s$, $+5^\circ 01' \cdot 8$ (1950). Minkowski (1957) states that the galaxy displays a double nucleus and rather strong emission lines of [O II] and [O III], indicating that an active collision may be in progress. The angular size of the galaxy is about $\frac{1}{2}'$ and that of the radio source about $2\frac{1}{2}'$ (Carter, unpublished data).

Finally, the coincidences noted in the catalogue between the source 08+010 and the Hydra II cluster of galaxies, and the source 23+02 and the Pegasus I cluster of galaxies are probably significant. The less good coincidence between the source 23+03 and the Pegasus II cluster may also be real. These and other observations of southern clusters will be discussed in a future paper.

IV. STATISTICS

Two properties of the distribution of radio sources in the catalogue area have been investigated statistically; these are the two-dimensional distribution across the celestial sphere and the distribution in depth based on counts to different flux density levels. Because of the finite resolution of the aerial, spatial clustering may affect the source counts substantially; clustering is therefore dealt with first, principally by consideration of the two-dimensional distribution and some observations of source sizes and background irregularities.

(a) *Clustering*

In paper I it was suggested that the radio sources might display large-scale clustering. This conclusion was based essentially on applying the χ^2 test to areas measuring 10° by 10° ; the distribution was found to be non-random at the 2 per cent. confidence level. We now have a much greater quantity of more homogeneous data, and similar tests which have been applied indicate that large-scale deviations from randomness are much less than before and perhaps not significant. The χ^2 test was applied to sources above various limiting intensity levels in areas of different sizes well away from the galactic circle. It is only with 30° by 30° squares that a significant non-randomness is indicated.

This applies to all sources, and also to sources stronger than $10^{-25} \text{ W m}^{-2} (\text{c/s})^{-1}$; in both cases, however, the confidence level is only 5 per cent. It appears that large-scale clustering, if present, is small; with the accumulation of more data it may be possible to usefully apply more sophisticated tests.

In paper I it was also noted that the number of "extended" sources listed in the catalogue was significantly more than expected from chance blending effects in a universe of randomly distributed "point" sources. This conclusion is strengthened by analysis of the present catalogue.

We have considered areas from $21^{\text{h}}-00^{\text{h}}-05^{\text{h}}$ and from $09^{\text{h}}-15^{\text{h}}$, that is, areas well away from the galactic circle. Of the sources with flux densities greater than $40 \times 10^{-26} \text{ W m}^{-2} (\text{c/s})^{-1}$, 20 are listed as "extended" and 2 as "perhaps extended": of sources stronger than $20 \times 10^{-26} \text{ W m}^{-2} (\text{c/s})^{-1}$, 36 are listed as "extended" and 17 as "perhaps extended". The numbers of chance blends classified as a single source which is "extended" or "perhaps extended" may be estimated as in paper I. The expected number of such blends having flux densities greater than $40 \times 10^{-26} \text{ W m}^{-2} (\text{c/s})^{-1}$ is 2, and the number greater than $20 \times 10^{-26} \text{ W m}^{-2} (\text{c/s})^{-1}$ is 15. These numbers are very much smaller than those observed. We therefore conclude that a substantial proportion of the stronger sources have an angular size which is resolvable with our aerial ($> \frac{1}{4}^\circ$) and/or that small-scale clustering effects are significant.

Both these explanations are consistent with present astronomical knowledge. It is now commonly accepted, for instance, that the majority of galaxies are organized into clusters of various sizes, and in a large cluster the conditions would seem favourable for the production of radio sources by collision or interaction; thus the existence of several physically related radio sources very close together is quite conceivable. There is also evidence for clustering of a higher order into "supergalaxies" (for example, de Vaucouleurs 1956) to which the same arguments apply. Clusters of galaxies in general emit much more at radio frequencies than the integrated emission of their component normal galaxies (e.g. Brown and Hazard 1953). This may be attributable to associated "radio galaxies" or sometimes equally well to radiation from the cluster as a whole, that is, to intergalactic emission. The latter process has been discussed in some detail by Shklovsky (1954) in its application to radiation from the local supergalaxy. Thus, on present evidence, it would seem quite possible that individual radio galaxies should appear relatively frequently in physically related groups of two or more, and that close clusterings of galaxies can create vast radio sources of large angular size even at very great distances. These possibilities will be considered quantitatively in future papers. A further possibility which should not be ignored is that some, at least, of the Class II sources of large angular size may be located in our galaxy. Analysis of variations in brightness of the background radiation observed with the 3.5 m cross aerial shows that some bright regions at moderate distances from the plane are probably related to the galactic emission: obvious examples in the present catalogue (not, however, included in the area analysed for obvious reasons) are the large sources associated with the nebulosities in Orion. Such aspects of the galactic emission will also be considered

in future papers. At the moment it is sufficient to note that a model comprising a random distribution of "point" sources is inadequate to explain all the observational data.

(b) *Source Counts*

In order to investigate the distribution of the radio sources in depth, source counts have been made to various limiting flux density levels. As in an earlier paper (Mills 1952) the sources have been divided into two classes, those within $12\frac{1}{2}^\circ$ of the galactic circle (Class I sources) and the remainder (Class II sources).

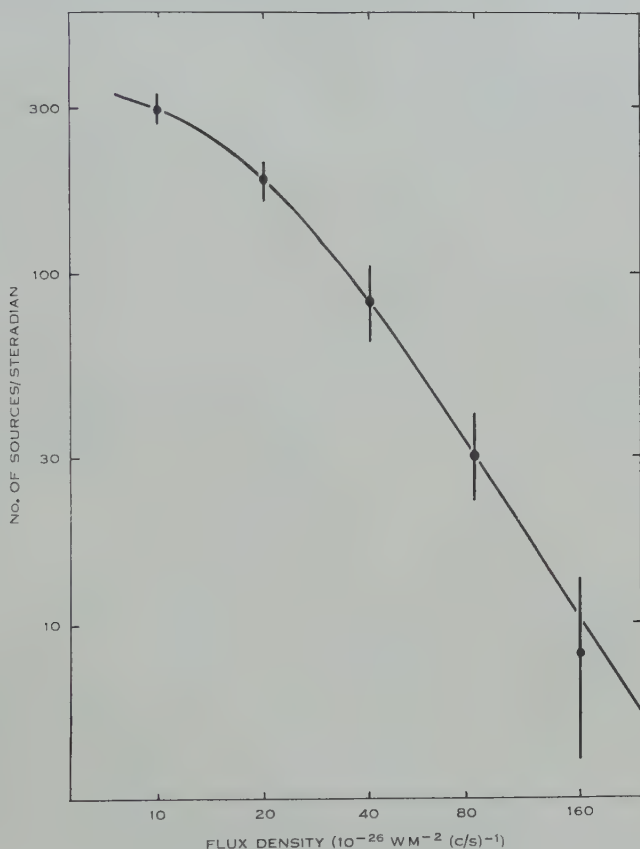


Fig. 1.—Counts of Class I sources, that is, sources within $12\frac{1}{2}^\circ$ of the galactic circle.

The counts are shown in Figures 1 and 2, where the logarithms of the number density of sources with flux densities S or higher are plotted against $\log S$; the standard errors in the plotted points due to chance effects in the distribution (\sqrt{N}) are shown as vertical wings in the figure. The actual numbers from which the diagrams were constructed are given in Tables 5 and 6.

There are clearly insufficient Class I sources for a detailed analysis. There are enough, however, to show that their statistics are greatly different from the Class II sources, which is consistent with the original results of Mills (1952); as

before, the smaller slope suggests an origin in the galactic disk for the majority of the stronger sources. It is interesting that there are substantial differences shown in the catalogue for the sources near the centre and anticentre of the Galaxy, those near the centre being, on the average, much stronger and more numerous. This again suggests a relation with the galactic structure, and it is clear that these sources must be considered in relation to the general distribution of the galactic emission. This will be done in some papers dealing with the Galaxy which are now in preparation.

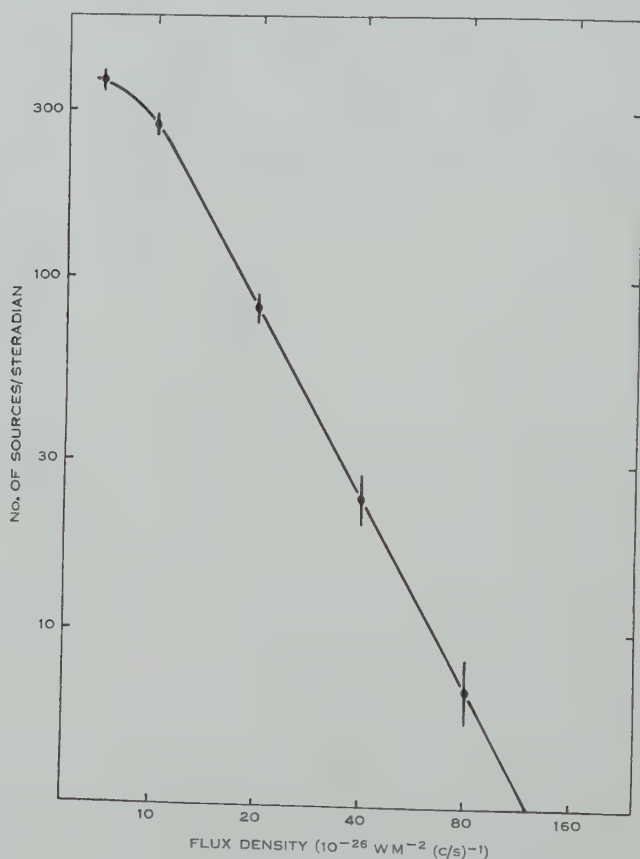


Fig. 2.—Counts of the Class II sources.

The Class II sources are sufficiently numerous for further analysis. The straight line shown in Figure 2 passing through the derived points has a slope of -1.8 . As discussed in paper I, there are two instrumental factors which contribute to this slope, the finite resolution of the aerial and the uncertainties in the weaker flux densities due to noise. However, with randomly distributed sources neither of the effects is large; from the data in paper I a mean increase of slope of about -0.15 is estimated, leaving a net slope of -1.65 . It is well known that, after all corrections, the slope should be -1.5 for a random distribution of "point" sources in a static Euclidean universe; thus there is, in addition

to the previous evidence of angular size and clustering, some evidence from the source counts that the above model is inadequate. However, the evidence is not conclusive, for inspection of Figure 2 shows that the standard errors due to statistical effects are large and the apparent small excess of faint sources could equally well be a statistical deficiency of close and strong sources.

It is interesting to compare the evidence from the angular size data and the source counts to see if (i) an excess of faint and distant sources, corresponding to a slope of -1.65 in the source counts ogive, can increase the number of blends sufficiently to account for the excessive number of "extended" sources, or (ii) a large number of "extended" sources can result in an increased apparent

TABLE 5
THE NUMBERS OF CLASS I SOURCES ABOVE DEFINED FLUX DENSITY LEVELS

Flux density (10^{-26} W m $^{-2}$ (c/s) $^{-1}$) ..	10	20	40	80	160
No. of sources, N (Total=156)	142	88	40	15	4

TABLE 6
THE NUMBERS OF CLASS II SOURCES ABOVE DEFINED FLUX DENSITY LEVELS

Flux density (10^{-26} W m $^{-2}$ (c/s) $^{-1}$) ..	7	10	20	40	80	160
No. of sources, N (Total=1003)	982	754	218	63	19	4

slope of the magnitude observed. It is easy to see from the earlier data that (i) is not possible; the increase in the number of blends would be negligible, as the parameters of the model source distribution were fixed largely by the numbers of faint sources. On the other hand it is evident that, if the clustering is greater than in a random distribution, the increased slope due to blending will be enhanced; quantitatively the attribution of all extended sources to blends is adequate to explain the observed slope. It might appear that sources of finite angular size could have no effect on the slope, since integrated fluxes are used in the counts. However, the possibility arises because the catalogue is restricted to sources of angular size less than 2° , in order to eliminate background irregularities as far as possible; it seems likely that, at the same time, many large-scale extragalactic concentrations are excluded. One obvious example which comes in this category is the "Local Supergalaxy".

In principle we may investigate the importance of these factors by comparing the proportion of extended sources listed at different flux density levels. The bottom end of the catalogue must be excluded from comparisons of this kind because of the difficulty of identifying an extended source: the lowest level at which consistent recognition seems possible is 40×10^{-26} W m $^{-2}$ (c/s) $^{-1}$. In Subsection (a) we gave the number of extended sources above this level as 20,

with 2 listed as "perhaps extended", in a defined area well away from the galactic circle. The total number of sources above the same level in the same area is 44; whence it would appear that about half the sources are either of large angular size or physical blends. For sources with flux densities greater than $160 \times 10^{-26} \text{ W m}^{-2} (\text{c/s})^{-1}$, that is, sources which, on the average, are at half the distance, the number of extended Class II sources listed in the whole southern sky is 3 out of a total 10 sources (excluding the Magellanic Clouds as belonging to the class of "normal" galaxies with which we are at present not concerned). Unfortunately these numbers are too low to permit a firm conclusion, although they are consistent with some systematic omission of very large sources. The data are obviously inadequate to correct for such effects in order to derive counts which are dependent only on the large-scale distribution in depth, which is required before they may usefully be applied to a cosmological model. Further information on these questions might be expected to come from the Cambridge interferometer surveys which discriminate strongly against large sources.

In paper I it was shown that the Cambridge 2C catalogue (Shakeshaft *et al.* 1955) can be accorded little weight, principally on account of the poor primary resolution of the instrument. Ryle (1956) has criticized this conclusion on the basis of conjecture as to the assumptions made in paper I, which was at that time unpublished. However, comparisons of the assumptions and method of analysis used in the paper with Ryle's conjectures shows that the latter were unfounded. The assessment of the reliability of 2C appears to have received confirmation in a report by Hewish (1957) of greatly improved agreement between Sydney and a more recent Cambridge survey. This survey has been made using their original instrument at double the 2C survey frequency and hence four times the resolution of the earlier work.

However, he also reports that the source counts again have a very large slope, -2.2 in one area and -2.7 in another, although somewhat less than the slope of -3.0 obtained in the original survey (Ryle and Scheuer 1955). We have no detailed information about these results but it would appear that the increased resolution has reduced the slope, and the question remains whether the excess slope is again the result of instrumental effects, or whether it is real. If the latter, it would suggest strongly, when taken in conjunction with our pencil-beam survey, that the effects of angular size or small-scale clustering are significant. Hewish also remarked that the statistics of the output envelope fluctuations are inconsistent with a uniform distribution of "point" sources. This has already been noted by Ryle and Scheuer in connexion with the 2C survey; but with the data supplied it was not possible to verify this, or to make use of their probability density distribution. It is to be hoped that some quantitative information will be published on this important point.

One of the regions discussed by Hewish is bounded by Right Ascensions 00^{h} and 08^{h} and declinations $+10^\circ$ and -10° ; that is, within our present catalogue. For this area he quotes a slope of -2.7 for the $\log N - \log S$ relation. We have performed a source count in this region and find a slope of -1.7 . It is interesting that if the *peak* values of the flux densities of our extended sources are used instead of the integrated fluxes, thus to some extent simulating the

results obtained with an interferometer, the slope is increased to -2.0 , making up at least some of the difference.

It is clear that the accumulation of data of this kind and intercomparisons between surveys carried out under different conditions will go a long way towards sorting out the complex picture. However, it is hoped that a more direct approach which is now being instituted will yield definite answers to the more important questions. Since the principal uncertainty is in the proportion of distant sources of small angular size, the most obvious procedure is to use an interferometer with sufficient spacing between aerials to respond to these alone. Such an instrument is now being put into operation at the Radiophysics Laboratory, in which the aerial spacing is 3000λ , yielding a lobe separation of 1.2 min of arc: the sensitivity is expected to be similar to that in our survey.

To conclude, it seems hardly necessary to point out the futility of attempting to analyse the cosmological implications of these source counts in detail until the above problems are sorted out. It would seem that, as in the optical case (although for different reasons), the straightforward counting of observed sources to various flux density levels is inadequate to define the form of the Universe. We have some hope that additional angular size data which we are planning to collect may make a significant contribution towards this end.

V. REFERENCES

- BEČVAR, A. (1951).—"Atlas Coeli Skalnate Pleso." (Přirodovedecké Vydavatelství: Prague.)
- BROWN, R. H., and HAZARD, C. (1952).—*Phil. Mag.* **43**: 137.
- BROWN, R. H., and HAZARD, C. (1953).—*Mon. Not. R. Astr. Soc.* **113**: 123.
- CARTER, A. W. L. (1955).—*Aust. J. Phys.* **8**: 564.
- HEWISH, A. (1957).—The distribution of radio stars. A paper presented at the American Astronomical Society Meeting, Urbana, Illinois.
- MILLS, B. Y. (1952).—*Aust. J. Sci. Res. A* **5**: 266.
- MILLS, B. Y. (1958).—Radio frequency radiation from external galaxies. "Handbuch der Physik." Vol. 53. (Springer-Verlag: Berlin.) (In press.)
- MILLS, B. Y., LITTLE, A. G., and SHERIDAN, K. V. (1956).—*Aust. J. Phys.* **9**: 218.
- MILLS, B. Y., LITTLE, A. G., SHERIDAN, K. V., and SLEE, O. B. (1958).—*Proc. Inst. Radio Engrs., N.Y.* **46**: 67.
- MILLS, B. Y., and SLEE, O. B. (1957).—*Aust. J. Phys.* **10**: 162.
- MINKOWSKI, R. (1957).—Radio Astronomy, Symposium No. IV of the International Astronomical Union, p. 107. (Cambridge Univ. Press.)
- RYLE, M. (1956).—*Sci. Amer.* **195** (6): 10.
- RYLE, M., and SCHEUER, P. A. G. (1955).—*Proc. Roy. Soc. A* **230**: 448.
- SHAKESHAFT, J. R., RYLE, M., BALDWIN, J. E., ELSMORE, B., and THOMPSON, J. H. (1955).—*Mon. Not. R. Astr. Soc.* **67**: 106.
- SHKLOVSKY, I. S. (1954).—*Astr. J. Moscow* **31**: 533.
- DE VAUCOULEURS, G. (1953).—A revision of the Harvard survey of bright galaxies. Aust. Nat. Univ. Mimeograph, Canberra.
- DE VAUCOULEURS, G. (1956).—"Vistas in Astronomy." (Ed. A. Beer.) Vol. 2, p. 1574. (Pergamon Press: London.)

ON THE RADIO EMISSION OF HYDROGEN NEBULAE

By C. M. WADE*

[*Manuscript received April 29, 1958*]

Summary

The interpretation of radio-frequency observations of H II regions is discussed with particular regard for the possible effects of random variations in the electron density and electron temperature through the nebulae. It is shown that such variations serve to alter the optical depth and that the conventional definition of the "emission measure" requires modification if it is to be considered an observable quantity. The radio emission of Strömgren spheres is discussed, and a means of determining their electron temperatures is described. An empirical method for the determination of Strömgren's constant defining the ionized volume as a function of the spectral type and luminosity of the exciting star is described.

I. INTRODUCTION

A number of galactic H II regions have been observed at radio frequencies. The observations published to date cover frequencies ranging from 19.7 to 9375 Mc/s. Such observations can provide information about the temperatures and densities of the nebulae, and about the far ultraviolet radiation of the stars exciting them. The objective of the present paper is to examine in some detail the problem of deriving physical data on the nebulae from radio observations.

The H II regions constitute a special class of radio source characterized by their spectra, which are "flat" except at the lower frequencies where they become optically thick. That is, their flux densities are nearly constant over a very wide range of frequencies. This is what is expected if the nebulae are radiating by the thermal process of free-free transitions in an ionized gas. They are readily distinguished from the "non-thermal" radio sources, whose spectra show a strong frequency dependence.

The radio emission of H II regions is well understood theoretically (e.g. Piddington 1951). Discussions of the nebulae as radio sources have generally assumed for simplicity that the objects are uniform throughout, although it has been known that a non-uniform distribution of the nebular gas would tend to increase the radio emission because the emissivity of an ionized gas depends on the square of the electron density. In the present paper we shall give particular attention to the consequences of such a non-uniform distribution.

Section II is devoted to a discussion of the directly observable properties of H II regions. Section III considers the effect of density and temperature variations on the optical depth, and the magnitude of the effect is estimated for some particular nebular models. A modification of the definition of the term

* Division of Radiophysics, C.S.I.R.O., University Grounds, Chippendale, N.S.W.

“ emission measure ” is suggested. In Section IV, the radio emission of Strömgren spheres is discussed, and a method of deriving their electron temperatures is described. Section V is concerned with the excitation of the nebulae.

II. DIRECTLY OBSERVABLE PROPERTIES

Mills, Little, and Sheridan (1956) have used the term “ apparent flux density ” to describe the observed radiation of an H II region of uniform electron temperature T_e , lying between the observer and a background of uniform brightness temperature T_b . This quantity is

$$F_{app} = (2kf^2/c^2)(T_e - T_b) \int (1 - e^{-\tau}) d\Omega, \quad \dots\dots\dots (1)$$

where the integration is extended over the entire solid angle Ω subtended by the nebula. In the above equation, k is Boltzmann's constant, c is the velocity of light, f is the frequency in hertz, and τ is the optical depth at the frequency f . The “ apparent flux density ” as defined here is the difference between the actual flux density coming from the nebula plus the transmitted portion of the background radiation, and the flux density which would be incident from the area covered by the nebula if the latter were absent. It should not be confused with the actual flux density coming from the area covered by the nebula, which is

$$(2kf^2/c^2) \{ T_e \int (1 - e^{-\tau}) d\Omega + T_b \int e^{-\tau} d\Omega \},$$

provided no emission arises between the nebula and the observer. This assumption is reasonable for frequencies above about 100 Mc/s for nebulae near enough to be observed optically. The quantity “ apparent flux density ” offers the advantage that it describes the “ visibility ” of a nebula in a convenient manner. If $T_e > T_b$, the object appears in emission; if $T_e < T_b$, it is seen in absorption. It will not be detectable at all if $T_e = T_b$. It is important to note that, while the actual flux density is always positive, the “ apparent flux density ” may be either positive or negative, depending on whether the object is seen in emission or absorption.

The apparent flux density incident from an H II region may be measured at various frequencies. Existing observations have been made with aërials having beamwidths of the same order as the angular sizes of the nebulae observed. In the future it will be possible, presumably, to resolve a number of the larger nebulae with very large aërials. This will permit direct determinations of the distributions of radio brightness across these objects. At present, however, the radio data usually consist only of measurements of F_{app} and T_b at various frequencies. Information about the angular sizes and apparent shapes of most of the nearer H II regions may be obtained from photographs, although in some cases these are severely affected by interstellar absorption. In the present paper we shall consider only the quantities derivable from these data.

Equation (1) implies that we may learn the electron temperature of a nebula simply by finding the frequency at which F_{app} is zero, and measuring T_b at this frequency. Then $T_e = T_b$. Mills, Little, and Sheridan (loc. cit.), following this line of thought, have shown that the commonly accepted value of $T_e \approx 10,000^\circ \text{K}$ is consistent with the observational evidence. There are two considerations

which prevent this approach being a sensitive method of determining electron temperatures. Firstly, current calibration methods do not permit the absolute accuracy of brightness temperature measurements to be better than about ± 20 per cent. Secondly, the random noise fluctuations in the output of contemporary receivers prevent an accurate determination of the frequency at which F_{app} vanishes. Therefore at present this method can give us only the order of magnitude of the electron temperatures. Nevertheless, future improvements in instrumental calibration techniques and receiver characteristics may make an approach of this kind useful.

III. THE RADIO EMISSION OF A NON-UNIFORM HYDROGEN NEBULA

(a) *The Optical Depth*

We shall consider the optical depth in a cloud of completely ionized hydrogen gas. We assume that the emission and absorption of radiation in the gas are completely described by the mechanism of free-free transitions. According to Piddington (1951), at low densities the absorption coefficient per centimetre in a gas consisting of equal numbers of protons and electrons is

$$\left. \begin{aligned} \kappa &= \zeta n_e^2 / f^2 T_e^{3/2}, \\ \zeta &= 9.70 \times 10^{-3} \ln(3kT_e/2hf). \end{aligned} \right\} \dots\dots\dots (2)$$

In these equations, n_e is the number of free electrons per unit volume and h is Planck's constant. The optical depth along a particular line of sight in a cloud of ionized hydrogen is

$$\tau = \int_0^S \kappa ds, \dots\dots\dots (3)$$

where S is the length traversed by the line in the ionized region. If the distribution of matter and temperature is perfectly uniform, we have simply

$$\tau_u = \zeta n_e^2 S / f^2 T_e^{3/2}. \dots\dots\dots (4)$$

The subscript u denotes the uniform case.

In general, there is no physical justification for an *a priori* assumption that an actual nebula is entirely uniform. The electron density and electron temperature at a point s on a particular line of sight may be written

$$\begin{aligned} n_e(s) &= \bar{n}_e + \delta n_e(s), \\ T_e(s) &= \bar{T}_e + \delta T_e(s), \end{aligned}$$

where \bar{n}_e and \bar{T}_e are the average values for the nebula, and $\delta n_e(s)$ and $\delta T_e(s)$ are the local deviations from the average. Defining

$$\begin{aligned} m &\equiv \delta n_e(s) / \bar{n}_e, \\ t &\equiv \delta T_e(s) / \bar{T}_e, \end{aligned}$$

we have

$$\left. \begin{aligned} n_e(s) &= \bar{n}_e(1+m), \\ T_e(s) &= \bar{T}_e(1+t). \end{aligned} \right\} \dots\dots\dots (5)$$

Neglecting the slow variation of ζ as a function of T_e , we may rewrite the equations (2) as

$$\left. \begin{aligned} x &= \frac{\zeta(\bar{n}_e)^2}{f^2 \bar{T}_e^{3/2}} \cdot \frac{(1+m)^2}{(1+t)^{3/2}}, \\ \zeta &= 9.70 \times 10^{-3} \ln(3k\bar{T}_e/2hf). \end{aligned} \right\} \dots\dots\dots (6)$$

The optical depth is then

$$\tau = \frac{\tau_u}{S} \int_0^S \frac{(1+m)^2}{(1+t)^{3/2}} ds. \dots\dots\dots (7)$$

The optical depth therefore differs from that in the uniform case by a factor

$$Q = \frac{1}{S} \int_0^S \frac{(1+m)^2}{(1+t)^{3/2}} ds. \dots\dots\dots (8)$$

We shall call this the "amplification factor". If $Q > 1$, the optical depth is greater than that corresponding to a uniform density and temperature; if $Q < 1$, it is less. The value of Q is independent of the frequency, and is fixed by the distribution of matter and temperature within the nebula. Equation (8) is quite general; it may be applied for either systematic or random variations.

It is evident that, in the absence of systematic density and temperature variations, the optical depth is the same as if the nebula were at a uniform temperature \bar{T}_e with a uniform density equal to $\bar{n}_e \sqrt{Q}$. An unfortunate consequence of these considerations is that the average density of an H II region cannot be found from radio-frequency observations without an independent knowledge of the distribution of matter and temperature. We may, however, derive an "equivalent density" defined by

$$n_{eq} = \bar{n}_e \sqrt{Q}. \dots\dots\dots (9)$$

This is "equivalent" in the sense that it is the density the nebula would need to have in order to produce the observed radio emission if it were uniform. Strictly speaking, it is a parameter which depends in an undetermined manner on the average electron density and the variations in density and temperature. In the special case that the electron temperature is uniform the equivalent density is equal to the root mean square density.

(b) The Apparent Flux Density

The apparent brightness temperature at a point on the projected surface of a nebula is

$$\begin{aligned} T_B &= e^{-\tau} \int_0^\tau T_e e^{\tau'} d\tau' + T_b e^{-\tau} \\ &= \bar{T}_e (1 - e^{-\tau}) + \bar{T}_b e^{-\tau} \int_0^\tau e^{\tau'} d\tau' + T_b e^{-\tau}. \end{aligned}$$

If the temperature fluctuations are random along the line of sight, and if their linear scale is small compared to the length of the line of sight in the ionized

region, $t=0$, and consequently the integral vanishes. Therefore, the apparent brightness temperature is

$$T_B = \overline{T_e}(1 - e^{-\tau}) + T_b e^{-\tau}.$$

This differs from T_b by

$$\Delta T = (\overline{T_e} - T_b)(1 - e^{-\tau}).$$

The apparent flux density is then

$$F_{app} = (2kf^2/c^2)(\overline{T_e} - T_b) \int (1 - e^{-\tau}) d\Omega. \quad \dots\dots\dots (10)$$

This expression is equivalent to equation (1) if the electron temperature is uniform.

(c) *On the Evaluation of the Amplification Factor*

There is no general solution to the equation for the amplification factor (equation (8)). A considerable simplification results, however, if the variations are purely random and if there is a functional relationship between electron temperature and density. In an actual nebula, it is likely that such a functional relationship exists, at least approximately. In the present subsection we shall consider two extreme possible cases: an isothermal nebula (uniform electron temperature) and an adiabatic nebula (uniform entropy). If the density variations within a nebula arise from turbulent motions of the gas, we may expect the actual relationship to lie somewhere between these two extremes.

In an isothermal nebula, $t=0$ everywhere and equation (8) gives

$$Q = \overline{(1+m)^2},$$

where the average is taken over the part of the line of sight lying in the nebula. Since we are assuming that the density variations are random, the above expression may be written

$$Q = 1 + \overline{m^2}.$$

In an adiabatic nebula, we have

$$1+t = (1+m)^{\gamma-1},$$

where γ is the ratio of specific heats. The electrons and protons comprising the gas have no communicable internal degrees of freedom, so $\gamma=5/3$. We obtain the result that

$$Q = \overline{1+m} = 1.$$

Thus the optical depth is unaffected by the presence of random density variations if they are adiabatically related to the temperature variations.

(d) *The Discrete Cloud Model*

The value of the amplification factor depends not only on the relation between m and t but also on the form of the function $m(s)$. A particularly simple model is one in which a fraction β of the mass of a nebula resides in condensations

occupying a fraction α of the nebular volume, the densities being uniform and equal in the condensations, and also uniform outside them. This situation corresponds to having

$$m = \beta/\alpha - 1$$

inside the condensations and

$$m = (\alpha - \beta)/(1 - \alpha)$$

outside. If the electron temperature is uniform, and if the condensations are distributed at random, we find

$$Q = \beta^2/\alpha + (1 - \beta)^2/(1 - \alpha). \quad \dots\dots\dots (11)$$

A consequence of this result is that

$$Q(\alpha, \beta) = Q(1 - \alpha, 1 - \beta).$$

This means that if, instead of condensations, we have uniform regions of sub-average density scattered through a nebula, we obtain a positive amplification factor equal to that in the inverse case. Figure 1 shows Q as a function of α and β

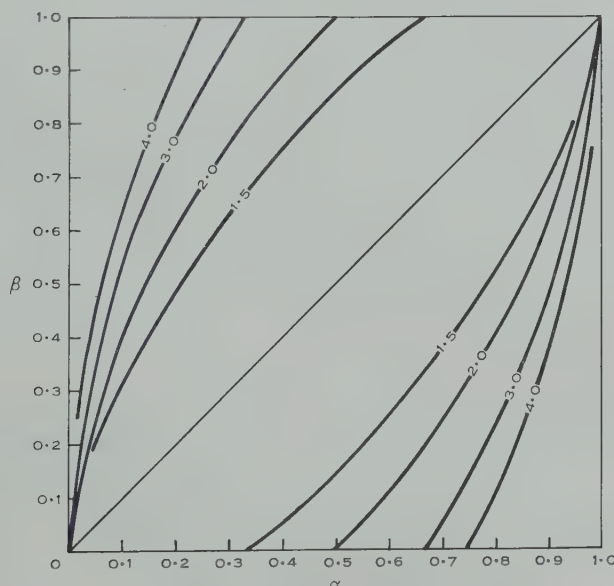


Fig. 1.— $Q(\alpha, \beta)$ for an isothermal nebula.

for constant electron temperature. The part of the diagram above the diagonal corresponding to $Q=1$ applies to the case of true condensations ($\alpha < \beta$); the part below refers to “negative condensations” ($\alpha > \beta$). If $\alpha = \beta$, the nebula is uniform and perforce $Q=1$.

(e) The Emission Measure

Strömgren (1948) introduced the term *emission measure* to describe the monochromatic intensity of nebular radiation in the Balmer lines. The emission

measure is defined as the product of the square of the density (which Strömgren assumed to be uniform) and the length of the emission path in parsecs ;

$$\epsilon = (3 \cdot 08 \times 10^{18})^{-1} n_e^2 S. \quad \dots\dots\dots (12)$$

The numerical factor is required because we are using c.g.s. units.

The emission measure has also been used by some authors in discussing the radio emission of H II regions. According to equation (4), we have a simple relationship between emission measure and optical depth for perfectly uniform nebulae :

$$\tau_u = 3 \cdot 08 \times 10^{18} \zeta_\epsilon / f^2 T_e^{3/2}. \quad \dots\dots\dots (13)$$

For a non-uniform nebula, however,

$$\tau = \zeta(\bar{n}_e)^2 S Q / f^2 \bar{T}_e^{3/2} = \zeta n_{eq}^2 S / f^2 \bar{T}_e^{3/2}. \quad \dots\dots\dots (14)$$

If we define

$$\epsilon' = (3 \cdot 08 \times 10^{18})^{-1} n_{eq}^2 S, \quad \dots\dots\dots (15)$$

we obtain an expression analogous to (13) :

$$\tau = 3 \cdot 08 \times 10^{18} \zeta \epsilon' / f^2 \bar{T}_e^{3/2}. \quad \dots\dots\dots (16)$$

The emission measure as defined by Strömgren is not appropriate for discussing a non-uniform nebula, since it cannot be determined from observations without reference to a physical model of the nebula. However, the analogous quantity ϵ' defined by equation (15) can be related directly to the optical depth, provided the average electron temperature is known. In the next section we shall outline a means of obtaining the average electron temperature and the equivalent density of a spherical nebula from radio data. Therefore ϵ' is in principle an observable quantity, whereas ϵ is not. We suggest that the term "emission measure" should be defined by (15) instead of (12). The two definitions are equivalent for a uniform nebula because then $Q=1$.

IV. STRÖMGREN SPHERES

Strömgren (1939, 1948) has studied theoretically the ionized region which would surround a hot star imbedded in an extended uniform cloud of hydrogen gas. He showed that the Lyman continuum radiation of the star would cause almost complete ionization of the hydrogen out to a quite sharply defined boundary. The radius of the ionized sphere was shown to depend upon the density of the gas and the Lyman continuum flux emitted by the star. Strömgren's theory provides a convenient basis for discussing hydrogen nebulae, although it treats a highly idealized case. Many galactic H II regions are very nearly spherical in shape and are reasonably concentric with the stars exciting them. Such nebulae evidently approximate to the case considered by Strömgren and are frequently referred to by the convenient designation "Strömgren spheres". In the present section we consider the radio emission of these objects.

(a) *The Apparent Flux Density*

We assume that the average density of the nebular hydrogen does not change as a function of distance from the exciting star and that the density and

temperature variations have a linear scale small compared to the radius of the ionized sphere. We shall let τ_0 denote the optical depth of the ionized sphere at its apparent centre, and θ_0 be the apparent angular radius of the nebula in radians. We also assume that the radius of the ionized region is small compared with its distance, so that we may set

$$\theta_0 = \sin \theta_0.$$

The optical depth at an angular distance θ from the centre of the nebula will be

$$\tau(\theta) = \tau_0(1 - \theta^2/\theta_0^2)^{\frac{1}{2}}.$$

For a spherical nebula, equation (10) takes the form

$$\begin{aligned} F_{app} &= 2kf^2/c^2(\overline{T_e} - T_b) \int_0^{\theta_0} (1 - e^{-\tau(\theta)}) 2\pi\theta d\theta \\ &= \frac{4\pi kf^2(\overline{T_e} - T_b)}{c^2} \int_0^{\theta_0} \theta(1 - e^{-\tau(\theta)}) d\theta. \end{aligned} \quad (17)$$

Completion of the integration gives the result

$$F_{app} = \{4\pi k\theta_0^2 f^2(\overline{T_e} - T_b)/c^2\} Y(\tau_0), \quad (18)$$

where

$$Y(\tau_0) = \frac{1}{2} + \tau_0^{-2} [e^{-\tau_0}(\tau_0 + 1) - 1]. \quad (19)$$

An equivalent expression is

$$Y(\tau_0) = \sum_{n=1}^{\infty} (-1)^{n+1} \frac{\tau_0^n}{n!(n+2)}. \quad (20)$$

This series may be shown to converge for all values of $\tau_0 > 0$. If $\tau_0 < 0.1$, the approximation

$$Y(\tau_0) = \frac{1}{3}\tau_0$$

is sufficiently accurate. We also note that $Y(\infty) = 0.500$. Figure 2 shows $Y(\tau_0)$ for $0.1 \leq \tau_0 \leq 10.0$.

Equation (18) takes the limiting forms

$$F_{app} = \{4\pi k\theta_0^2 f^2(\overline{T_e} - T_b)/3c^2\} \tau_0, \quad (21)$$

for $\tau_0 \ll 1$ and

$$F_{app} = 2\pi k\theta_0^2 f^2(\overline{T_e} - T_b)/c^2, \quad (22)$$

for $\tau_0 \gg 1$. These correspond to expressions given by Mills, Little, and Sheridan (loc. cit.).

(b) The Electron Temperature

Equation (18) provides a means of estimating the average electron temperature of a Strömgren sphere from observations of the apparent flux density at two well-separated frequencies. The central optical depth as a function of electron temperature may be calculated at each frequency. The ratio of the optical depths at each temperature is then found. The ratio corresponding to

the correct electron temperature will be equal to the ratio of the opacities at the two frequencies, so

$$\frac{\tau_{0,1}}{\tau_{0,2}} = \frac{\zeta_1}{\zeta_2} \cdot \frac{f_2^2}{f_1^2} \dots\dots\dots (23)$$

where the subscripts 1 and 2 denote the two frequencies.

We may illustrate the calculation by considering the Rosette Nebula (NGC 2237). This object resembles a classical Strömgren sphere except for the fact that its central part appears to be much lower in density than the average. According to Mills, Slee, and Hill (1958), the apparent flux density at 85.5 Mc/s is $2.7 \times 10^{-24} \text{ W m}^{-2} \text{ Hz}^{-1}$. Piddington and Trent (1956) found the apparent

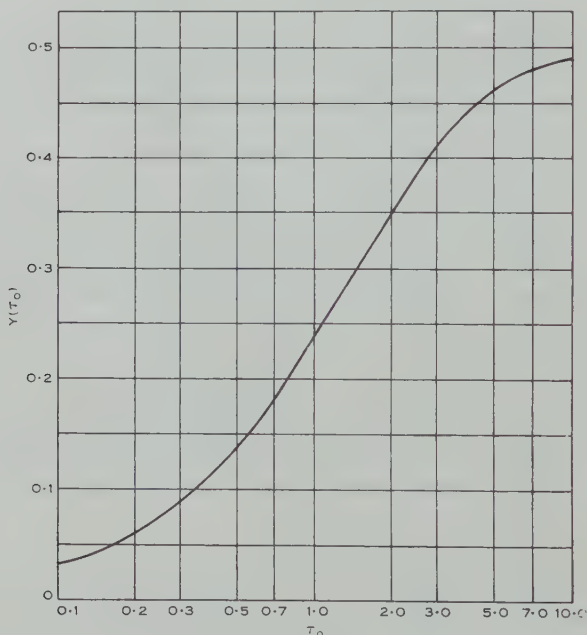


Fig. 2.—The function $Y(\tau_0)$.

flux density at 600 Mc/s to be $4.0 \times 10^{-24} \text{ W m}^{-2} \text{ Hz}^{-1}$. At 85.5 Mc/s, $T_b = 1800^\circ \text{K}$ (Mills, unpublished data), while T_b is negligible at 600 Mc/s. The angular diameter of the nebula is 80 min of arc. Using these data, we find the values given in Table 1. The values of the ratio of $\tau_{0,1}$ to $\tau_{0,2}$ calculated from equation (23) are 56.6 at 8000°K and 56.4 at $10,000^\circ \text{K}$. We can now plot the two determinations of optical depth ratio as a function of electron temperature. The intersection of the two curves gives the required value of the average electron temperature. We do this in Figure 3, obtaining the result $\bar{T}_e = 8600^\circ \text{K}$. Unfortunately, this result is sensitive to the errors in the measured flux densities; it is probably within about 30 per cent. of the correct value. The method has other limitations. Firstly, it is not reliable if the central optical depth is too

great at one or both frequencies, since the computed values of τ_0 are then too sensitive to errors in $Y(\tau_0)$. Secondly, if the nebula is too thin optically at both frequencies, so that $Y(\tau_0)$ is linear at each, one cannot obtain a reliable result.

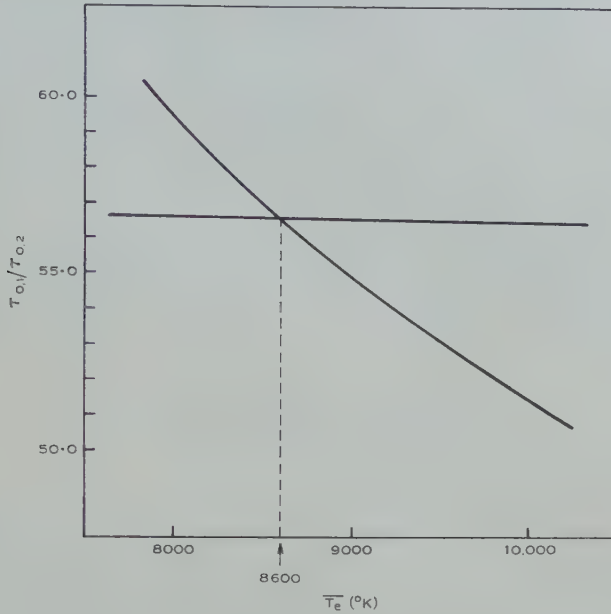


Fig. 3.—Determination of the electron temperature of the Rosette Nebula (NGC 2237).

(c) *The Equivalent Density*

We may use equation (18) to derive the equivalent density of a spherical nebula from measurements of its radio-frequency flux, provided its central optical depth is not too great. If F_{app} , \bar{T}_e , T_b , and θ_0 are known we may solve for $Y(\tau_0)$, and find the corresponding value of τ_0 .

Now

$$\tau_0 = 6 \cdot 16 \times 10^{18} \theta_0 R \zeta n_{eq}^2 / f^2 \bar{T}_e^{3/2},$$

whence

$$n_{eq} = 5 \cdot 72 \times 10^{-10} f (\tau_0 \bar{T}_e^{3/2} / 2 \zeta R \theta_0)^{1/2}. \quad \dots\dots\dots (24)$$

TABLE 1

CENTRAL OPTICAL DEPTH OF THE ROSETTE NEBULA AS A FUNCTION OF FREQUENCY AND ELECTRON TEMPERATURE

Frequency	85.5 Mc/s		600 Mc/s		
\bar{T}_e ($^{\circ}K$)	$Y(\tau_0, 1)$	$(\tau_0, 1)$	$Y(\tau_0, 2)$	$\tau_0, 2$	$\tau_0, 1/\tau_0, 2$
8,000	0.229	0.96	0.00535	0.0161	59.5
9,000	0.197	0.78	0.00475	0.0142	55.0
10,000	0.173	0.66	0.00428	0.0128	51.5

R is the distance to the nebula in parsecs, which must be found from optical studies of the exciting star. This ordinarily requires a knowledge of the absolute magnitude and intrinsic colour of the star. These data are not well determined for O, B, and Wolf-Rayet stars. Since only stars of these types are sufficiently hot and luminous to excite an observable Strömgren sphere, it follows that the distance to an H II region is not easy to find accurately. The seriousness of the matter is reduced somewhat for our purposes by the fact that the equivalent density depends only on the square root of R . We note that the derived equivalent density is almost independent of \bar{T}_e , since the product $f^2 \bar{T}_e^{3/2} \tau_0$ is nearly constant for a given nebula over a very wide range of temperatures.

Johnson (1957) has found that the distance to the cluster NGC 2244, which contains the stars exciting the Rosette Nebula, is 1660 parsecs. Using this distance and the radio data given above, we obtain an equivalent density of 17 cm^{-3} , in good agreement with the value 14 cm^{-3} found by Minkowski (1955) from optical data.

V. THE EXCITATION OF THE NEBULAE

Strömgren (1939) has shown that the quantity

$$U = n_{eq}^{2/3} s_0 \dots\dots\dots (25)$$

is a constant depending on the spectral type and absolute luminosity of the exciting star, s_0 being the radius of the ionized zone in parsecs. Strömgren assumed a uniform density in his derivation; according to the considerations we have presented in Section III this may be replaced by the equivalent density n_{eq} .

Strömgren computed the excitation constant U for stars of various spectral types from his theory. The calculated values depend on a number of assumptions, however, and it is desirable to have a direct observational determination of these quantities. Substituting (24) in (25) and replacing $R\theta_0$ by s_0 , we get

$$U = 6.87 \times 10^{-7} s_0^{2/3} f^{2/3} \bar{T}_e^{1/2} (\tau_0 / 2\zeta)^{1/3} \dots\dots\dots (26)$$

This result is independent of the possible presence of density variations in the nebular gas, since τ_0 varies as s_0^{-2} for any value of Q . It is almost independent of the electron temperature, because of the quasi-constancy of the product $f^2 \bar{T}_e^{3/2} \tau_0$. Therefore we may find U for a star exciting a spherical nebula if we know s_0 from optical studies and τ_0 from radio observations.

We may apply equation (26) to the Rosette Nebula, using the data of Section IV. We find $U = 126$. The nebula is actually excited by four O-stars—one each of types O5 and O6, and two of type O8. The effective U for a group of stars is the cube root of the sum of the cubes of the U 's of the individual stars. The effective U found from the data given by Strömgren is 167, which is in fair agreement with the empirical value.

VI. ACKNOWLEDGMENT

The research described in the present paper was made possible by a Junior Research Fellowship granted by the Commonwealth Scientific and Industrial Research Organization.

VII. REFERENCES

- JOHNSON, H. L. (1957).—*Astrophys. J.* **126** : 121.
- MILLS, B. Y., LITTLE, A. G., and SHERIDAN, K. V. (1956).—*Aust. J. Phys.* **9** : 218.
- MILLS, B. Y., SLEE, O. B., and HILL, E. R. (1958).—*Aust. J. Phys.* **11** : 360.
- MINKOWSKI, R. (1955).—Gas dynamics of cosmic clouds, I.A.U. Symposium Ser. No. 2, p. 3.
(North Holland Publishing Co.: Amsterdam.)
- PIDDINGTON, J. H. (1951).—*Mon. Not. R. Astr. Soc.* **111** : 45.
- PIDDINGTON, J. H., and TRENT, G. H. (1956).—*Aust. J. Phys.* **9** : 74.
- STRÖMGREN, B. (1939).—*Astrophys. J.* **89** : 526.
- STRÖMGREN, B. (1948).—*Astrophys. J.* **108** : 242.

AN INVESTIGATION OF THE STRONG RADIO SOURCES IN CENTAURUS, FORNAX, AND PUPPIS

By K. V. SHERIDAN*

[*Manuscript received May 20, 1958*]

Summary

Isophotes of three strong southern radio sources have been prepared from observations at a wavelength of 3.5 m with a 50 min pencil-beam system. The two extragalactic sources Centaurus-A and Fornax-A are found to have extensive coronas very much larger than the size of the associated galaxies. Using optical estimates of distance these two sources are shown to have similar spatial extensions and may be physically similar systems. The third source, Puppis-A, may be the remnants of a galactic supernova of type II.

I. INTRODUCTION

The three strongest southern radio sources, Centaurus-A, Fornax-A, and Puppis-A, have been reliably identified with visible objects, the first two with the external galaxies NGC 5128 and NGC 1316 respectively (Bolton, Stanley, and Slee 1949; Mills 1952*b*, 1954) and the last with a peculiar galactic nebulosity (Baade and Minkowski 1954*a*). They have all been shown to have an angular extent of the order of 1° or more, and in one case (Centaurus-A) it is known that the object comprises at least two distinct emitting components (Mills 1953; Bolton *et al.* 1954). The major part of the published material is based on interferometer observations which had to be interpreted under various simplifying assumptions.

The aim of the present paper is to describe new observations of these sources with a high-resolution, pencil-beam system producing records which can be readily and directly interpreted. Contour diagrams of the brightness distributions across the sources have been prepared using the Sydney Mills Cross at a wavelength of 3.5 m. This instrument has a beamwidth between half-power response points of about 50 min of arc, and, while this is insufficient to resolve the fine structure of the sources, it is adequate to obtain a general picture of their size and shape. When used in conjunction with the interferometer data much useful information may be obtained.

The most striking feature of the observations is the extremely large dimensions of the two extragalactic sources, considerably more than that of the visible portions of the galaxies. Taking estimates of distance based on optical data it appears that the spatial extent may be similar in each example.

* Division of Radiophysics, C.S.I.R.O., University Grounds, Chippendale, N.S.W.

II. OBSERVATIONS AND REDUCTIONS

The radio telescope and its use have been described in detail by Mills *et al.* (1958). In the present case it is operated as a meridian transit instrument, and the aerial beam is automatically switched successively to five declinations separated by approximately 20 min of arc; the beam is held at each declination for a period of 12 sec. The central position of the beams is set by manual adjustments to the aerial. A typical record showing the passage of the Centaurus

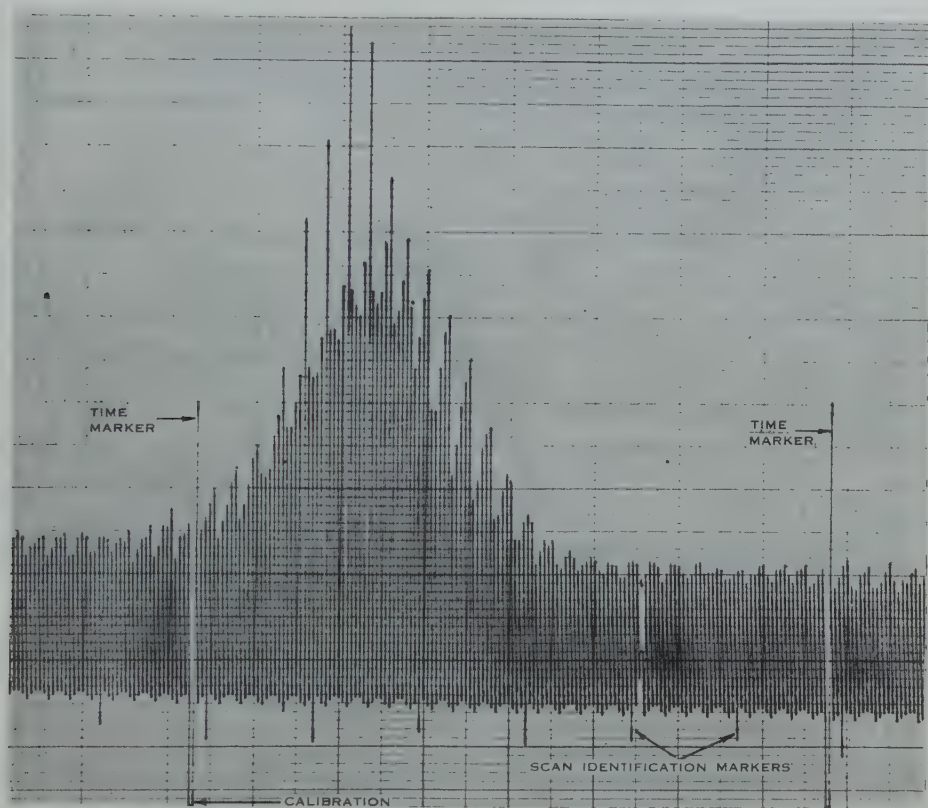


Fig. 1.—A record of a passage of Centaurus-A through the scanning aerial beam. Five separate declinations are recorded quasi-simultaneously.

source through the aerial beam is shown in Figure 1. The recorder deflection at the top of one of the lines is proportional to the average aerial temperature observed with the aerial pointed to one declination for the 12-sec period. The five declinations are scanned from south to north so that every fifth line on the chart corresponds to the same declination. The aerial temperatures* are therefore observed at points on a grid with spacings 1 min in Right Ascension

* The procedure necessary to correct for various instrumental effects is described by Mills *et al.* (1958). These corrections, which include those for beam asymmetries and collimation error, have been applied to the data quoted subsequently.

TABLE I
3·5 M RADIO MEASUREMENTS AND ASSOCIATED OPTICAL DATA

Radio Source	I.A.U. No.	Position (epoch 1950)*		Flux Density* S (10^{-24} W m $^{-2}$ (c/s) $^{-1}$)	Estimated Angular Size†		Radio Mag- nitude‡ m_R	Associated Optical Objects				$m_R - m_P$	Approx. Dis- tance § (kpc)
		R.A.	Dec.		Major Axis	Minor Axis		NGC No.	R.A.	Dec.	Photo. Mag- nitude m_P		
Centaurus-A	13S4A	13 ^h 22 ^m . 4 \pm 0 ^m . 2	-42° 41' \pm 4'	87 \pm 13	6°	2°	1·7	5128	13 ^h 22 ^m . 47	-42° 45' . 6	6·5	-4·8	750
Fornax-A	03S3A	03 ^h 20 ^m . 6 \pm 0 ^m . 2	-37° 23' \pm 3'	9·5 \pm 1·5	0·7°	\approx 0·5°	4·2	1316	03 ^h 20 ^m . 7	-37° 25'	9·5	-5·3	5000
Puppis-A	08S4A	08 ^h 20 ^m . 9 \pm 0 ^m . 3	-42° 52' \pm 4'	6·9 \pm 1	0·8°	0·8°	4·5	Galactic nebu- losity	08 ^h 20 ^m . 3	-42° 48'	—	—	0·5

* Errors shown are estimated probable errors based on uncertainties in system parameters.

† Angular size estimates for Fornax-A and Puppis-A derived on assumption of Gaussian distribution across the sources.

‡ Radio magnitude defined by $m_R(3\cdot5m) = -53\cdot4 - 2\cdot5 \log_{10} S(3\cdot5m)$.

§ Distance estimates for Fornax-A and Centaurus-A (de Vaucouleurs 1956) and Puppis-A (Baade and Minkowski 1954a).

and 20 min of arc in declination. These spacings are sufficiently small to define the distribution uniquely within the resolution limits of the aerial (Bracewell 1956).

The records include the effects of both source and background radiation. The background temperature can be removed by interpolating between regions around the source. This process was easily applied to the Fornax-A records, as the background level is practically uniform in this region, but was rather more difficult in the case of the Puppis-A source, which is superimposed on a steep galactic gradient.

The removal of the background was not performed on the Centaurus-A source because it covers such a large area that interpolation becomes very subjective; consequently it was decided to present the direct observational data.

Observations of the Fornax and Puppis regions required records at 15 declinations, involving three different manual settings of the aerial. However, the five most important sections straddling each source were obtained within periods of about 10 min. Because of the great size of the Centaurus source records at over 30 declinations were required. To minimize the effects of any calibration changes occurring in the period during which the records were obtained, the background temperatures were matched at a particular Right Ascension (viz. $13^{\text{h}} 01^{\text{m}}$) where the temperature was nearly uniform over the declination range concerned. After smoothing the temperatures at this Right Ascension (to discriminate against the possible inclusion of small sources) and plotting against declination, the calibration changes were detected by noting if groups of five consecutive points were displaced to one side of the curve. Ten sections were corrected, one group by 10 per cent. and the others by about 5 per cent. This process merely smoothed the background level, to avoid irregularities in the final plot.

For each source an additional set of curves was constructed along hour circles (using the above east-west sections) and used to interpolate in declination. To complete these curves, use was made of the interpolation theorem of Bracewell and Roberts (1954) to compute extra points at declination positions between the east-west sections. The values so obtained are those which would result if the aerial beam was actually pointed to these positions.

Both sets of curves were used to construct contour maps for each source, following the method outlined by Mills *et al.* (1958). The procedure was straightforward for Fornax-A and Puppis-A but more complicated for Centaurus-A owing to the effects of aerial side lobes in the north-south directions in which Centaurus-A is greatly extended. These small effects were removed by a smoothing process.

III. DISCUSSION

The results of these observations are summarized in Table 1 which sets out the radio features together with some optical data and estimates. The three sources will be discussed in turn.

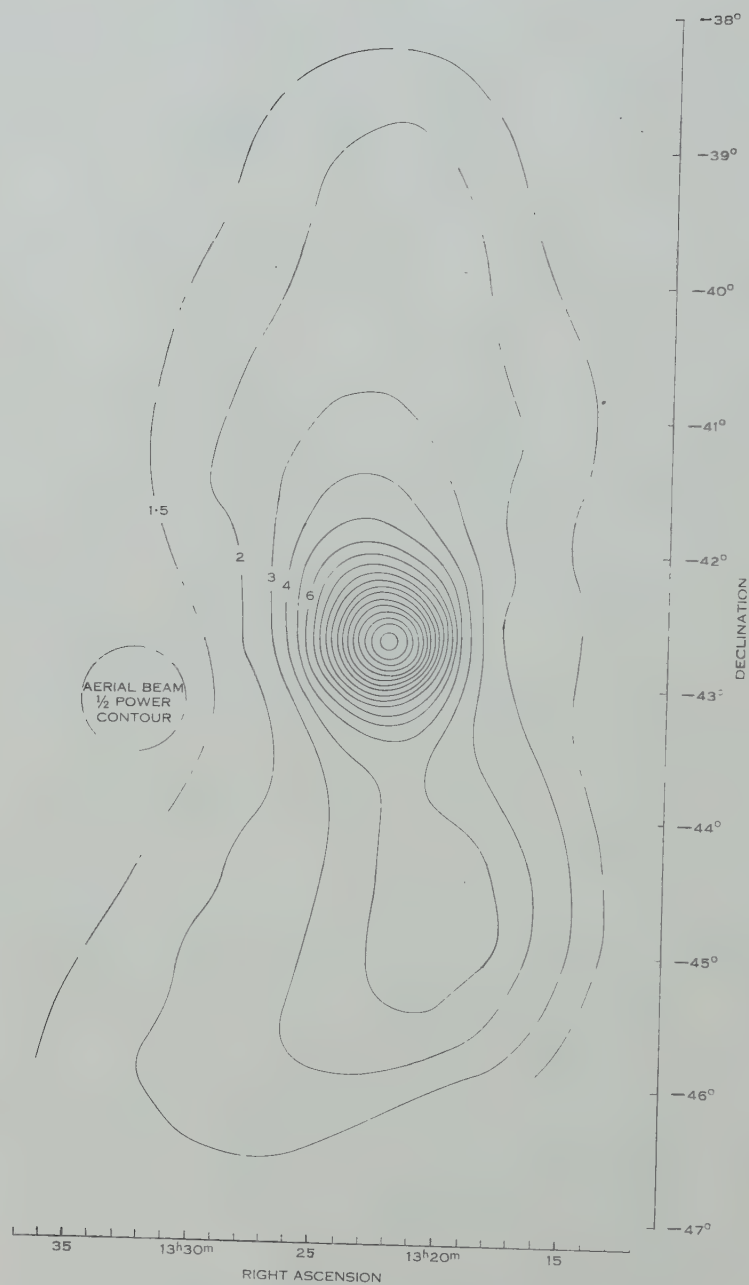


Fig. 2.—Radio isophotes of Centaurus-A at a wavelength of 3.5 m (epoch 1950), contour interval 3000 °K.

(a) Centaurus-A

Radio emission from this source was first detected by Bolton (1948) using a 3 m wavelength sea interferometer. He reported its angular width as less than 15 min of arc. Subsequent observations have revealed two contributions: intense radiation from a small area of about 6 by 3' coincident with the position of the galaxy NGC 5128, and a much more extensive and less bright area surrounding the galaxy. Mills (1953) measured the brightness distribution across the source with an interferometer at a wavelength of 3 m and found about 45 per cent. of the total energy in the central concentration, with a maximum brightness temperature of about 4×10^6 °K.

A preliminary investigation to determine the main features of the radio brightness distribution was made with the present aerial in an early state of its development and the result has been compared with the optical brightness distribution (de Vaucouleurs and Sheridan 1957). The radio results were presented as contours of apparent equal brightness temperature above a smooth interpolated reference level but were uncorrected for aerial beam asymmetries and side lobes. The main result of this comparison was that a reasonable agreement existed in the east-west directions but the radio distribution was much more extensive in the north-south directions.

Figure 2 shows radio isophotes (epoch 1950) constructed from the later, more accurate, observations described above. The contour interval is 3000 °K and the galactic background has not been removed. The background distorts the outermost contours slightly, but otherwise will have little effect on the distribution. The contours show a large source, of apparent maximum temperature about 5×10^4 °K, considerably elongated in position angle about 12°. Emission is intense over an area of about 2 by 6° and is strongly concentrated in the central region with the peak at $13^h 22^m 4 \pm 0^m 2$, $-42^\circ 41' \pm 4'$. This peak corresponds closely to the "point source" at the centre which is, of course, not resolved. To determine the integrated flux density S of the source the background was first subtracted by smoothly interpolating through the base of each section and then evaluating, over the remaining contours, the expression $S = (2k/\lambda^2) \int T d\Omega$. The value obtained was $S = (87 \pm 13) \times 10^{-24}$ W m⁻² (c/s)⁻¹. The uncertainty in this value (± 15 per cent.) is an estimated probable error based on uncertainties in the system parameters. Approximately 25 per cent. of this flux density is contained in the central concentration.

The extended emission between declinations -35° to -46° has a shape bearing slight resemblance to the letter "S". Piddington and Trent (1956) suggested a "link" between this source and the Galaxy (at a wavelength of 0.5 m) but no evidence in support of this was found in the 3.5 m results.

Extensive radio coronas are perhaps associated with most galaxies, a fact consistent with an origin in the emission of high energy electrons by the synchrotron mechanism, but the Centaurus-A source is peculiar in having such an elongated shape. Baade and Minkowski (1954*b*) consider that the interaction of two separate galaxies is involved. The distance of the system is not definitely established, but it has been estimated by de Vaucouleurs (1956) to be about

750 kpc. At this distance the linear dimensions of the extended source would be about 90 by 30 kpc. At the distance adopted by Burbidge and Burbidge (1957), 2.5 Mpc, the linear dimensions would be greater by a factor of more than 3. This seems unlikely, and their estimates of total emission from the source are probably rather too high.

(b) *Fornax-A*

Discovered in 1948 by Stanley and Slee, this source has since been identified by Mills (1954) with the galaxy NGC 1316. From preliminary observations with the Cross, Mills described it as an elongated object of overall size about 1° with the major axis in position angle about 160° .

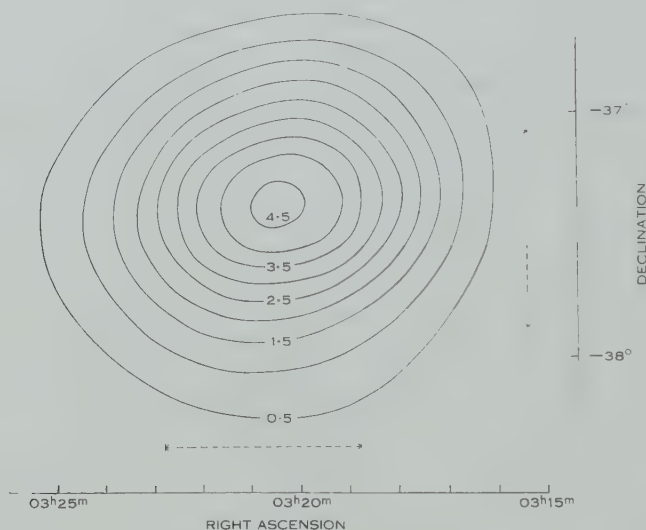


Fig. 3.—Radio isophotes of Fornax-A at a wavelength of 3.5 m (epoch 1950), contour interval 1500°K . The dashed lines indicate the aerial beamwidth between half-power points in the directions shown.

Figure 3 shows the recently determined contours, at 3.5 m, free from background emission, the contour interval being 1500°K . The centroid position of this emission is at $03^{\text{h}} 20^{\text{m}}.6 \pm 0^{\text{m}}.2$, $-37^\circ 23' \pm 3'$ (epoch 1950), a value which agrees with Mills's determination, and is almost exactly the position of the galaxy.

By integration over the contours of Figure 3 the flux density, at a wavelength of 3.5 m, is $S = (9.5 \pm 1.5) \times 10^{-24} \text{ W m}^{-2} (\text{c/s})^{-1}$.

This source is only barely resolved by the aerial and therefore the somewhat irregular shape of the contours can only indicate that the source itself possesses some structure. Its effective angular size* is estimated to be about 0.7 by 0.5° . The orientation of the contours appears to change with distance from the centre

* To obtain an effective angular size θ_e , a Gaussian distribution was assumed for the source and use made of the expression $\theta_e = (\theta_0^2 - \theta_b^2)^{1/2}$, where θ_0 = angle between half-power contours of the observed distribution, θ_b = angular width of the aerial beam between half-power points.

but the overall effect suggests a position angle of about 115° . A photograph shown by Baade and Minkowski (1954*c*, p. 130) indicates the position angle of the main body of the nebula as about 45° , and reveals some absorption patches along a line whose position angle is about 135° . These figures suggest that the strong radio emission might be associated in some way with the appearance of the absorption patches, but little weight can be given to this opinion until radio observations with higher resolving power are available.

The radio emission extends over a much greater area than the galaxy. Taking the distance as 5 Mpc (de Vaucouleurs 1956) and the effective angular size as about 0.7 by 0.5° the radio corona has a linear extent of about 60 by 40 kpc. This is of a similar order of magnitude to that of Centaurus-A,

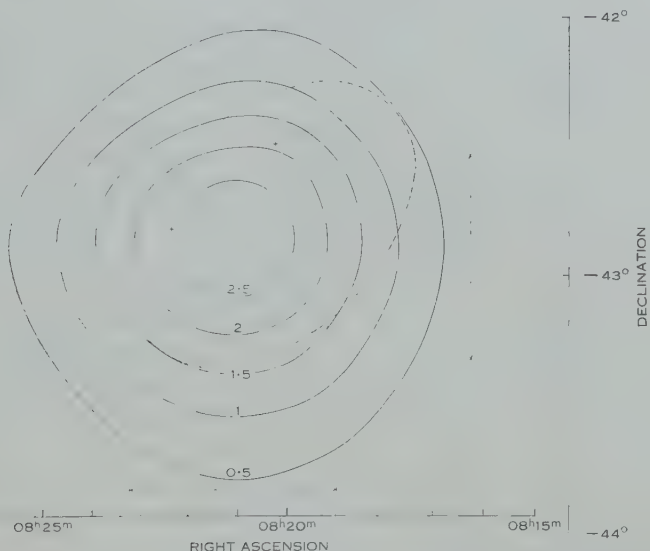


Fig. 4.—Radio isophotes of Puppis-A at a wavelength of 3.5 m (epoch 1950), contour interval 1500 °K. The dotted ellipse defines the boundary of the filamentary nebulosity associated with this source. The dashed straight lines indicate the aerial beamwidth between half-power points in the directions shown.

suggesting that the two sources might have similar physical origins. It is not yet possible to decide whether there is any similarity in the detailed distribution of the radiation in the two sources, in particular whether Fornax-A exhibits a structure combining an intense localized source and diffuse corona as suggested by Burbidge and Burbidge (1957).

(c) *Puppis-A*

The radio source, Puppis-A, was discovered by Stanley and Slee (1950). Using the position they quoted together with an angular size determined by Mills (1952*a*), the source was identified by Baade and Minkowski (1954*a*) with a network of gaseous filaments similar to those in Cassiopeia.

Figure 4 shows the radio isophotes, at a wavelength of 3.5 m, for this source. The background emission has been removed from these

contours, which have intervals of 1500°K . The centroid position of the source is at $08^{\text{h}} 20^{\text{m}} \cdot 9 \pm 0^{\text{m}} \cdot 3$, $-42^{\circ} 52' \pm 4'$ (epoch 1950), and the integrated flux density is $(6 \cdot 9 \pm 1) \times 10^{-24} \text{ W m}^{-2} (\text{c/s})^{-1}$. The source is definitely resolved and the effective angular size is estimated to be about $0 \cdot 8^{\circ}$.

Included in Figure 4 is a dotted ellipse which marks the boundary of the nebulosity discovered by Baade and Minkowski (1954*a*). Its coincidence in position with the radio source is very close. Emission from the radio source extends over a slightly greater area than the optical source. Although there is some uncertainty in the shapes of the contours owing to possible errors in the removal of the galactic background, it is possible that the slight bulging of the contours on the northern and eastern edges is associated with two of the brightest parts of the nebulosity whose approximate positions are indicated by the crosses in Figure 4. The interferometer observations of Mills (1952*a*) indicate that there is no outstanding region of emission of small angular size associated with Puppis-A. Because of its similarity to Cassiopeia, Puppis-A may be a supernova of type II. The weaker radio emission from Puppis-A may indicate that it is an older supernova.

An estimate of the distance to this source using hydrogen line absorption techniques is very desirable to determine the spatial extent and total emission.

IV. ACKNOWLEDGMENTS

The author wishes to express his gratitude to Mr. B. Y. Mills for his help and guidance during all phases of this work. He is also indebted to Mr. C. A. Shain for valuable assistance and criticism, to Mr. A. G. Little for contributing data on the aerial system, and to Mr. J. P. Wild for helpful suggestions regarding the manuscript.

V. REFERENCES

- BAADE, W., and MINKOWSKI, R. (1954*a*).—*Astrophys. J.* **119**: 206.
 BAADE, W., and MINKOWSKI, R. (1954*b*).—*Astrophys. J.* **119**: 215.
 BAADE, W., and MINKOWSKI, R. (1954*c*).—*Observatory* **74**: 130.
 BRACEWELL, R. N. (1956).—*Aust. J. Phys.* **9**: 297.
 BRACEWELL, R. N., and ROBERTS, J. A. (1954).—*Aust. J. Phys.* **7**: 615.
 BOLTON, J. G. (1948).—*Nature* **162**: 141.
 BOLTON, J. G., STANLEY, G. J., and SLEE, O. B. (1949).—*Nature* **164**: 101.
 BOLTON, J. G., WESTFOLD, K. C., STANLEY, G. J., and SLEE, O. B. (1954).—*Aust. J. Phys.* **7**: 96.
 BURBIDGE, G. R., and BURBIDGE, E. MARGARET (1957).—*Astrophys. J.* **125**: 1.
 MILLS, B. Y. (1952*a*).—*Aust. J. Sci. Res. A* **5**: 266.
 MILLS, B. Y. (1952*b*).—*Aust. J. Sci. Res. A* **5**: 456.
 MILLS, B. Y. (1953).—*Aust. J. Phys.* **6**: 452.
 MILLS, B. Y. (1954).—*Observatory* **74**: 248.
 MILLS, B. Y., LITTLE, A. G., SHERIDAN, K. V., and SLEE, O. B. (1958).—*Proc. Inst. Radio Engrs.*, N.Y. **46**: 67.
 PIDDINGTON, J. H., and TRENT, G. H. (1956).—*Aust. J. Phys.* **9**: 74.
 STANLEY, G. J., and SLEE, O. B. (1950).—*Aust. J. Sci. Res. A* **3**: 234.
 DE VAUCOULEURS, G. (1956).—*Occ. Notes R. Astr. Soc.* **18**: 118.
 DE VAUCOULEURS, G., and SHERIDAN, K. V. (1957).—Radio Astronomy Symposium No. IV of the International Astronomical Union, p. 169. (Cambridge Univ. Press.)

ON THE CYLINDRICAL PROBE METHOD OF MEASURING THERMAL CONDUCTIVITY WITH SPECIAL REFERENCE TO SOILS

II. ANALYSIS OF MOISTURE EFFECTS

By D. A. DE VRIES* and A. J. PECK*

[*Manuscript received February 10, 1958*]

Summary

Thermal conductivity measurements in unsaturated moist porous media are complicated by the action of gravity on the moisture and by moisture movement caused by temperature gradients. The effects of these factors are discussed for the cylindrical probe method.

A theoretical analysis shows that temperature gradients give rise to a decrease of moisture content in the vicinity of the probe and a (much smaller) increase at greater radial distances from the probe axis. It is shown how these moisture changes depend on the physical properties of the medium, the probe radius, and the time of heating. The magnitude of the decrease of moisture content at the probe surface is approximately inversely proportional to the probe radius; it increases with increasing time and temperature. In our experiments the region of reduced moisture content has a thickness of the order of the probe radius (0.05 cm). The absolute change in moisture content and the influence on the measured value of the conductivity are small at temperatures below about 40 °C.

Gravity causes a decrease of moisture content with height in a manner determined by the relation between moisture content and water pressure. In probe experiments the effect of this variation of moisture content in the vertical direction is negligible at water pressures below about —100 cm (using atmospheric pressure as a datum).

Some experimental results are presented and analysed. The importance of the cooling branch in detecting the influence of moisture movement is illustrated.

I. INTRODUCTION

In Part I (de Vries and Peck 1958) we treated the measurement of the thermal conductivity of soils and thermal insulating materials by means of cylindrical probes with special reference to the probe characteristics. It was assumed for simplicity that the medium surrounding the probe could be characterized by an overall thermal conductivity and an overall thermal diffusivity that were independent of place and time.

The use of overall or "macroscopic" thermal properties is justified when we are dealing with dimensions that are large in comparison with the scale of inhomogeneity in the medium (e.g. pore and grain sizes). The assumption of their constancy with place and time can only hold approximately in unsaturated porous media, because in such materials moisture distributes itself unevenly under the influence of gravity and temperature gradients. In the present paper we discuss the effect of these phenomena on thermal conductivity measurements by means of the cylindrical probe method.

* Division of Plant Industry, C.S.I.R.O., Deniliquin, N.S.W.

After explaining the symbolism in Section II, a recently developed theory of simultaneous transfer of heat and moisture in unsaturated porous media is outlined in Section III. The theory of moisture movement in porous materials under the influence of a temperature gradient was discussed previously by one of us in collaboration with J. R. Philip (Philip and de Vries 1957). In this paper differential equations governing the simultaneous transfer of heat and moisture were presented. In a later paper (de Vries 1958) the theory was extended; differential equations of greater generality were derived (equations (9) and (10) below) and applied to the problem of steady-state heat conduction in moist porous media. We refer to the original papers for a detailed discussion of these phenomena and for further references.

In Section IV the (linearized) equations for heat and moisture transfer are solved for radial flow and the solution is applied to the cylindrical probe method. The influence of gravity had to be neglected in order to preserve radial symmetry. Gravity effects are discussed separately in Section V. Experimental data are presented and discussed in Section VI.

II. NOTATION AND UNITS

- a , Thermal diffusivity ($\text{cm}^2 \text{sec}^{-1}$),
- b , constant defined by equation (17) ($^\circ\text{C}$),
- c_l , specific heat of liquid water ($\text{cal g}^{-1} \text{ }^\circ\text{C}^{-1}$),
- C , volumetric heat capacity ($\text{cal cm}^{-3} \text{ }^\circ\text{C}^{-1}$),
- D_{atm} , molecular diffusion coefficient of water vapour in air ($\text{cm}^2 \text{sec}^{-1}$)
($=4.42 \times 10^{-4} \text{T}^{2.3}/P$),*
- $D_T = D_{Tl} + D_{Tv}$, thermal moisture diffusivity ($\text{cm}^2 \text{sec}^{-1} \text{ }^\circ\text{C}^{-1}$),
- $D_{Tl} = K \partial \Psi / \partial T$, thermal liquid diffusivity† ($\text{cm}^2 \text{sec}^{-1} \text{ }^\circ\text{C}^{-1}$),
- $D_{Tv} = f D_{\text{atm}} v (d\rho_0/dT) h(\nabla T)_a / \rho_l \nabla T$, with $f = S$ for $\theta_l < \theta_{lK}$, $f = (S - \theta_l) / [1 + \theta_l / (S - \theta_{lK})]$
for $\theta_l > \theta_{lK}$, thermal vapour diffusivity ($\text{cm}^2 \text{sec}^{-1} \text{ }^\circ\text{C}^{-1}$),
- $D_\theta = D_{\theta l} + D_{\theta v}$, isothermal moisture diffusivity ($\text{cm}^2 \text{sec}^{-1}$),
- $D_{\theta l} = K \partial \Psi / \partial \theta_l$, isothermal liquid diffusivity ($\text{cm}^2 \text{sec}^{-1}$),
- $D_{\theta v} = \alpha a D_{\text{atm}} v g \rho_v (\partial \Psi / \partial \theta_l) / \rho_l R_w T$, isothermal vapour diffusivity ($\text{cm}^2 \text{sec}^{-1}$),
- g , acceleration due to gravity (cm sec^{-2}),
- h , relative humidity,
- j , mechanical equivalent of heat (erg cal^{-1}),
- \mathbf{k} , unit vector in vertical direction,
- K , unsaturated hydraulic conductivity (cm sec^{-1}),
- L , heat of vaporization of water (cal g^{-1}),
- p , partial pressure of water vapour (mm Hg),
- P , total gas pressure (mm Hg),
- \mathbf{q}_h , heat flux density ($\text{cal cm}^{-2} \text{sec}^{-1}$),
- \mathbf{q}_l , liquid flux density ($\text{g cm}^{-2} \text{sec}^{-1}$),
- $\mathbf{q}_m = \mathbf{q}_l + \mathbf{q}_v$, moisture flux density ($\text{g cm}^{-2} \text{sec}^{-1}$),
- \mathbf{q}_v , vapour flux density ($\text{g cm}^{-2} \text{sec}^{-1}$),

* See Philip and de Vries (1957) for a discussion of the experimental values of D_{atm} .

† The value of the temperature coefficient of Ψ has been taken as that of the temperature coefficient of surface tension; see Philip and de Vries (1957) for a discussion of this point.

- Q , heat production per unit length of probe ($\text{cal cm}^{-1} \text{sec}^{-1}$),
 r , radial distance from axis of probe (cm),
 R , probe radius (cm),
 R_w , gas constant of water vapour ($\text{erg g}^{-1} \text{°C}^{-1}$),
 S , porosity (cm^3/cm^3),
 t , time (sec),
 T , absolute temperature (°K),
 T_0 , reference temperature, also initial temperature (°K),
 u , auxiliary variable (°K),
 z , vertical coordinate, positive upwards (cm),
 α , tortuosity factor for diffusion of gases in soil (≈ 0.67),
 β , ratio of volumetric heat capacity of probe to volumetric heat capacity of soil,
 γ , Euler's constant ($= 0.5772$),
 ϵ , auxiliary variable defined by equation (31),
 $\zeta = r/R$,
 $\theta = \theta_l + \theta_v$, total volumetric moisture content,
 θ_l , volumetric liquid content,
 θ_{lK} , value of θ_l at which "liquid continuity" fails,
 θ_v , volumetric vapour content (cm^3 of precipitable water vapour/ cm^3),
 κ , auxiliary diffusivity defined by equation (19) ($\text{cm}^2 \text{sec}^{-1}$),
 λ , thermal conductivity ($\text{cal cm}^{-1} \text{sec}^{-1} \text{°C}^{-1}$),
 λ' , auxiliary thermal conductivity defined by equation (23),
 $\nu = P/(P-p)$, mass flow factor,
 ρ_l , density of liquid water (g cm^{-3}),
 ρ_v , density of water vapour (g cm^{-3}),
 ρ_0 , density of saturated water vapour (g cm^{-3}),
 $\sigma = 4\kappa t/R^2$,
 $\tau = 4at/R^2$,
 $\Phi = \Psi + z$, moisture potential (cm),
 Ψ , water pressure with atmospheric pressure as datum (cm),
 $(\nabla T)_a$, average temperature gradient in air-filled pores (°C cm^{-1}).

III. THE EQUATIONS FOR SIMULTANEOUS TRANSFER OF HEAT AND MOISTURE

Heat and moisture transfer in unsaturated moist porous media are very complex phenomena. Moisture moves in the liquid phase under the influence of gravity and pressure gradients, and in the vapour phase under the influence of gradients of vapour density. Heat is transferred by conduction, radiation, and convection (including sensible heat transfer by liquid and vapour movement) and also as latent heat by vapour movement.

One of the principal difficulties encountered in developing a mathematical-physical theory of these phenomena arises from the intricate nature of the internal geometry of a porous material. This difficulty can be circumvented to a certain extent by the introduction of "macroscopic" parameters, which are proper averages taken over a sufficiently large volume.

Examples of such macroscopic parameters are the thermal conductivity λ and the hydraulic conductivity K , which occur in the equations for the heat and liquid flux densities, namely,

$$\mathbf{q}_h = -\lambda \nabla T, \quad \dots \dots \dots (1)$$

and

$$\mathbf{q}_l/\rho_l = -K \nabla \Phi = -K \nabla \Psi - K \mathbf{k}, \quad \dots \dots \dots (2)$$

In equation (1) it is assumed, of course, that heat transfer is by conduction only. Equation (2) is a mathematical expression of Darcy's law (see e.g. Baver 1956). Ψ is negative in unsaturated porous media; both Ψ and K are functions of the moisture content.

For the vapour flux density we have

$$\mathbf{q}_v = -D_v \nabla \rho_v, \quad \dots \dots \dots (3)$$

where D_v is a macroscopic diffusion coefficient.

The moisture contents in the liquid (θ_l) and vapour (θ_v) phases are not independent, of course. When there is equilibrium between both phases in the pores these quantities are related by the expression

$$\theta_v = (S - \theta_l)(\rho_0/\rho_l) \exp(g\Psi/R_w T). \quad \dots \dots \dots (4)$$

Both Ψ and ρ_v are functions of θ_l and T . The equations (2) and (3) can therefore be written as

$$\mathbf{q}_l/\rho_l = -D_{\theta_l} \nabla \theta_l - D_{Tl} \nabla T - K \mathbf{k}, \quad \dots \dots \dots (5)$$

$$\mathbf{q}_v/\rho_l = -D_{\theta_v} \nabla \theta_l - D_{Tv} \nabla T, \quad \dots \dots \dots (6)$$

from which the total moisture flux density \mathbf{q}_m follows by addition

$$\mathbf{q}_m/\rho_l = -D_\theta \nabla \theta_l - D_T \nabla T - K \mathbf{k}. \quad \dots \dots \dots (7)$$

The various diffusivities occurring in these equations are explained in the previous section. For their derivation we refer to the paper by Philip and de Vries (1957).

The equation for the heat flux density becomes

$$\mathbf{q}_h = -(\lambda - L\rho_l D_{Tv}) \nabla T + L\mathbf{q}_v + c_l(T - T_0)\mathbf{q}_m, \quad \dots \dots \dots (8)$$

where T_0 is an arbitrary reference temperature. The effect of heat transfer by vapour movement under the influence of temperature gradients is best included in λ , because it can be the dominant factor in the heat transfer in the gas-filled pores. A more complete discussion of this point and of the approximate character of (8) (which is due to limitations inherent in the "macroscopic" approach) is given by de Vries (1958).

By application of the conservation laws for energy and mass the following simultaneous differential equations in θ_l and T are obtained (de Vries 1958):

$$\left(1 + \frac{D_{\theta_v}}{\alpha v D_{\text{atm}}} - \frac{\rho_v}{\rho_l}\right) \frac{\partial \theta_l}{\partial t} + \frac{(S - \theta_l)h}{\rho_l} \frac{d\rho_0}{dT} \frac{\partial T}{\partial t} = \nabla(D_\theta \nabla \theta_l) + \nabla(D_T \nabla T) + \frac{\partial K}{\partial z}, \quad \dots \dots \dots (9)$$

$$\left[\frac{L\rho_l D_{\theta_v}}{\alpha v D_{\text{atm}}} - L\rho_v + j^{-1}g\rho_l\left(\Psi - T \frac{\partial \Psi}{\partial T}\right)\right] \frac{\partial \theta_l}{\partial t} + \left[C + L(S - \theta_l)h \frac{d\rho_0}{dT}\right] \frac{\partial T}{\partial t} \\ = \nabla(\lambda \nabla T) + L\rho_l \nabla(D_{\theta_v} \nabla \theta_l) + \rho_l c_l [(D_{\theta_l} \nabla \theta_l + D_{Tl} \nabla T + K \mathbf{k}) \nabla T]. \quad \dots (10)$$

It must be noted that all coefficients in (9) and (10) are functions of the dependent variables θ_i and T .

For completeness the assumptions underlying the present theory are listed below ; we refer to the original papers (Philip and de Vries 1957 ; de Vries 1958) for their discussion.

1. Liquid flow can be described by Darcy's law, which implies *inter alia* that it is laminar.
2. A continuous equilibrium exists between vapour and liquid contents per unit of volume, as expressed by equation (4).
3. The water pressure Ψ is a unique function of θ_i .
4. The "macroscopic" equation (8) holds to a good approximation.

As applied to soils these assumptions present no serious limitations with the exception of No. 3, which excludes cases where hysteresis operates. A discussion of this problem has been published by Philip (1955, 1957a).

In addition, the following approximations were made in the derivation of (9) and (10) :

- (a) The relative humidity is independent of temperature at constant θ_i .
- (b) Heat transfer by convection and radiation, viscous generation of heat, and transfer of sensible heat by water vapour are negligible.

These approximations are entirely justified in the applications to be considered here.

In the case of thermal conductivity experiments using cylindrical probes equations (9) and (10) must be solved for the region $r > R$ (the probe radius) with the following initial and boundary conditions :

$$T = T_0, \quad \Phi = \Phi_0, \quad \text{for } t = 0 \text{ and for } r = \infty \text{ (} t = \text{finite)}, \quad \dots \quad (11)$$

$$q_h = Qf(t)/2\pi R \approx Q/2\pi R, \quad \text{for } r = R \text{ and } t > 0, \quad \dots \dots \dots (12)$$

$$q_m = 0, \quad \text{for } r = R. \quad \dots \dots \dots (13)$$

Here $f(t)$ depends on the characteristics of the probe and the medium. Neglecting the influence of moisture movement, we have (see equations (7), (19), and (21) of Part I) :

$$f(t) = 1 - \beta\tau^{-1} + O(\tau^{-2}). \quad \dots \dots \dots (14)$$

For the present purpose we shall take $f(t)$ equal to unity ; the resulting error will be small in comparison with those arising from other approximations.

The equations can be solved in principle by the application of numerical methods. However, such a procedure promises to be very laborious. We shall therefore follow a different path, i.e. by studying an analytical solution for the approximation of constant coefficients. Although this means a drastic simplification it is believed that the results obtained in this way retain the essentials of the exact solution. This point will be discussed further in Section IV (b).

IV. APPROXIMATE ANALYTICAL SOLUTION

(a) *Theory*

An analytical solution can be obtained for simultaneous diffusion equations with constant coefficients, viz.:

$$C_{11}\partial\theta_l/\partial t + C_{12}\partial T/\partial t = D_{11}\nabla^2\theta_l + D_{12}\nabla^2T, \quad \dots\dots\dots (15)$$

$$C_{21}\partial\theta_l/\partial t + C_{22}\partial T/\partial t = D_{21}\nabla^2\theta_l + D_{22}\nabla^2T. \quad \dots\dots\dots (16)$$

Equations (9) and (10) reduce to this form when transfer of sensible heat is neglected and when all parameters (except θ_l and T) can be considered as constants. To preserve radial symmetry we shall further neglect the influence of gravity in this subsection, i.e. the term $K\mathbf{k}$ in equation (7).

Equations (15) and (16) can be solved by the introduction of a new variable (Crank 1956)*

$$u = T + b\theta_l, \quad \dots\dots\dots (17)$$

It can be easily verified that, when b is a root of

$$(C_{11} - bC_{12})(D_{21} - bD_{22}) = (C_{21} - bC_{22})(D_{11} - bD_{12}), \quad \dots\dots (18)$$

equations (15) and (16) reduce to simple diffusion equations:

$$\partial u/\partial t = \kappa \nabla^2 u, \quad \dots\dots\dots (19)$$

where

$$\kappa = \{C_{21}D_{12} - C_{11}D_{22} + b(C_{12}D_{22} - C_{22}D_{12})\} / (C_{12}C_{21} - C_{11}C_{22}). \quad \dots\dots (20)$$

Two different values of κ (and hence two transformed equations (19)) are obtained by substituting the two roots of (18) in (20).

Combining (17) with (11) and (12), (13) leads to the following boundary conditions for u :

$$u = u_0, \text{ for } t=0 \text{ and for } r=\infty \text{ (} t=\text{finite)}, \quad \dots\dots (21)$$

$$-2\pi R\lambda' \partial u/\partial r = Q, \text{ for } r=R, \quad \dots\dots\dots (22)$$

with

$$\lambda' = (\lambda D_\theta - L\rho_l D_{\theta v} D_T) / (D_\theta - bD_T). \quad \dots\dots\dots (23)$$

The solution to equation (19) subject to conditions (21) and (22) is (Carslaw and Jaeger 1950)

$$u = -\frac{Q}{\pi^2 \lambda' R} \int_0^\infty (1 - e^{-\kappa x^2}) \frac{J_0(xr)Y_1(xR) - Y_0(xr)J_1(xR)}{x^2[J_1^2(xR) + Y_1^2(xR)]} dx, \quad \dots\dots (24)$$

where J and Y are Bessel functions of the first and second kind. Numerical values for $r=R$ are also given by Carslaw and Jaeger (loc. cit., p. 283).

* We are indebted to Mr. J. R. Philip for drawing our attention to this solution.

In subsequent developments both "small time" and "large time" solutions are needed. The former can be found by the method described by Carslaw and Jaeger (loc. cit., p. 274). We find

$$\begin{aligned} \frac{2\pi\lambda'}{Q}(u-u_0) = & \zeta^{-\frac{1}{2}} \left[\sigma^{\frac{1}{2}} \operatorname{erfc} \frac{\zeta-1}{\sigma^{\frac{1}{2}}} - \frac{(1+3\zeta)\sigma}{8\zeta} i^2 \operatorname{erfc} \frac{\zeta-1}{\sigma^{\frac{1}{2}}} \right. \\ & \left. + \frac{(9+6\zeta+33\zeta^2)\sigma^{3/2}}{128\zeta^2} i^3 \operatorname{erfc} \frac{\zeta-1}{\sigma^{\frac{1}{2}}} + 0 \left(\sigma^2 i^4 \operatorname{erfc} \frac{\zeta-1}{\sigma^{\frac{1}{2}}} \right) \right], \end{aligned}$$

..... (25)

with $\zeta=r/R \geq 1$ and $\sigma=4\lambda t/R^2 < 1$.

The "large time" solution calculated by the method of Part I is

$$\frac{4\pi\lambda'}{Q}(u-u_0) = \ln \sigma - 2 \ln \zeta - \gamma + \frac{2}{\sigma} (\ln \sigma - \gamma) + \frac{1}{\sigma} (1 - 2 \ln \zeta + \zeta^2) + 0(\sigma^{-2}),$$

..... (26)

for $\sigma \gg 1$. This solution is analogous to equation (23) of Part I; the two are identical for $r=R$ and $\beta=0$.

(b) Numerical Examples

Numerical values of the coefficients occurring in equations (9) and (10) can be calculated when Ψ , K , λ , and C are known as functions of θ_i and T . Values of these quantities and the various moisture diffusivities for Yolo light clay were published by Philip (1957b) and for a medium sand (particle size between 0.01 and 0.06 cm) by de Vries (1958).

When a temperature gradient is set up in a soil that is initially in a state of equilibrium ($T=T_0$, $\Phi=\Phi_0$) water is caused to move in a direction opposite to that of the temperature gradient. This movement is counteracted by a return flow due to the resulting gradients of Φ . The moisture flux due to temperature gradients is proportional to D_T , the flux due to moisture potential gradients to D_Φ . In a stationary state with $q_m=0$ these two fluxes balance (see de Vries (1958) for a discussion of steady-state heat conduction).

In the probe experiments to be considered below the temperature gradient is of the order of 1 °C/cm or less and hence the moisture gradient (expressed as cm^{-1}) will be of the order of magnitude of D_T/D_Φ (expressed as $^\circ\text{C}^{-1}$) or less. Values of D_T/D_Φ at 20 °C for the two soils mentioned above are shown in Figure 1. It will be noted that moisture gradients must be very small at both high and low moisture contents. They can only be detectable in an intermediate moisture range corresponding to the peak in the D_T/D_Φ curve. The width and position of this range depend *inter alia* on the soil texture.

A consideration of the magnitudes of the various coefficients in equations (15) and (16) shows that at these intermediate moisture contents the following approximate values of b , κ , and λ' hold.

$$\begin{aligned} b_1 &= -a/D_T, & \kappa_1 &= D_\Phi, & \lambda'_1 &= CD_\Phi, \\ b_2 &= L\rho_t D_{\Phi v}/\lambda, & \kappa_2 &= a, & \lambda'_2 &= \lambda. \end{aligned}$$

This leads to the approximate solutions :

$$u_1 - u_{1,0} = T - T_0 - a(\theta_l - \theta_{l0})/D_T = \frac{Qt^{\frac{1}{2}}}{\pi C(D_0 R r)^{\frac{1}{2}}} \left[\operatorname{ierfc} \frac{r-R}{2D_0^{\frac{1}{2}}t^{\frac{1}{2}}} + \theta \left(\sigma^{\frac{1}{2}} \operatorname{ierfc} \frac{\zeta-1}{\sigma^{\frac{1}{2}}} \right) \right], \quad \dots \dots \dots (27)$$

$$u_2 - u_{2,0} = T - T_0 + L\rho_l D_{0v}(\theta_l - \theta_{l0})/\lambda = (Q/4\pi\lambda) [\ln 4at/r^2 - \gamma + \theta(\tau^{-1})], \quad \dots (28)$$

where θ_{l0} is the initial value of θ_l . Since $b_1 \gg b_2$ we have

$$\theta_l - \theta_{l0} = (D_T/a)[(u_2 - u_{2,0}) - (u_1 - u_{1,0})]. \quad \dots \dots \dots (29)$$

This solution yields negative values of $\theta_l - \theta_{l0}$ for small values of $r - R$, since here $u_1 - u_{1,0}$ is large in comparison with $u_2 - u_{2,0}$. However, the former quantity decreases much more rapidly with increasing r than the latter so that $\theta_l - \theta_{l0}$ is positive beyond a certain distance from the probe.

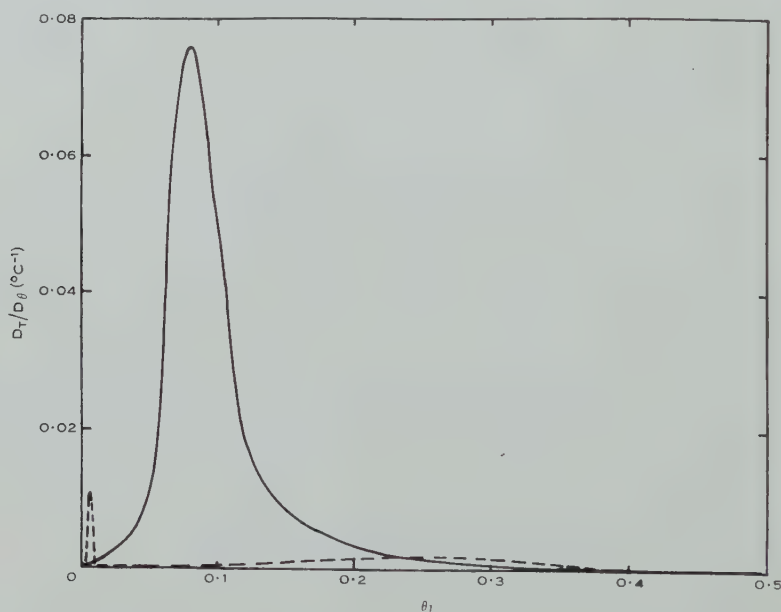


Fig. 1.—Values of D_T/D_0 in relation to volumetric moisture content, θ_l . Full curve for Yolo light clay (porosity $S=0.495$), broken curve for medium sand ($S=0.424$).

As a numerical example we consider the case of Yolo light clay at 20 °C and with $\theta_l=0.082$, which corresponds to the maximum of D_T/D_0 in Figure 1. Then the following values of the coefficients in equations (15) and (16) are found :

$$\begin{aligned} C_{11} &= 1.000, & C_{12} &= 4.16 \times 10^{-7} \text{ } ^\circ\text{C}^{-1}, & D_{11} &= 3.17 \times 10^{-6} \text{ cm}^2 \text{ sec}^{-1}, \\ D_{12} &= 2.42 \times 10^{-7} \text{ cm}^2 \text{ sec}^{-1} \text{ } ^\circ\text{C}^{-1}, & C_{21} &= -2.09 \text{ cal cm}^{-3}, \\ C_{22} &= C = 0.316 \text{ cal cm}^{-3} \text{ } ^\circ\text{C}^{-1}, & D_{21} &= 9.30 \times 10^{-4} \text{ cal cm}^{-1} \text{ sec}^{-1}, \\ D_{22} &= \lambda = 1.65 \times 10^{-3} \text{ cal cm}^{-1} \text{ sec}^{-1} \text{ } ^\circ\text{C}^{-1}. \end{aligned}$$

Substitution in equations (18), (20), and (23) leads to :

$$\begin{aligned} b_1 &= 2.18 \times 10^4 \text{ }^\circ\text{C}, \quad b_2 = 0.568 \text{ }^\circ\text{C}, \quad \kappa_1 = 3.02 \times 10^{-6} \text{ cm}^2 \text{ sec}^{-1}, \\ \kappa_2 &= a = 5.22 \times 10^{-3} \text{ cm}^2 \text{ sec}^{-1}, \quad \lambda'_1 = 0.948 \times 10^{-6} \text{ cal cm}^{-1} \text{ sec}^{-1} \text{ }^\circ\text{C}^{-1}, \\ \lambda'_2 &= \lambda = 1.65 \times 10^{-3} \text{ cal cm}^{-1} \text{ sec}^{-1} \text{ }^\circ\text{C}^{-1}. \end{aligned}$$

In our experiments we use probes with a radius of approximately 5×10^{-2} cm. The temperature rise is measured during a heating period of 180 sec or less, the first reading being taken at about 10 sec (cf. Part I). The change in moisture content, as calculated from equation (25) for $u_1 - u_{1,0}$ and equation (26) for $u_2 - u_{2,0}$ with the above numerical values of the parameters and $R = 0.05$ cm, $Q/4\pi\lambda = 0.1 \text{ }^\circ\text{C}$, is shown in Figure 2 for $t = 10$ and 180 sec. The increase in θ_l at some distance from the probe is so small as to be hardly discernible in the figure. The curve for $t = 180$ sec, for instance, has a maximum of about 2×10^{-5} at $r = 0.18$ cm.

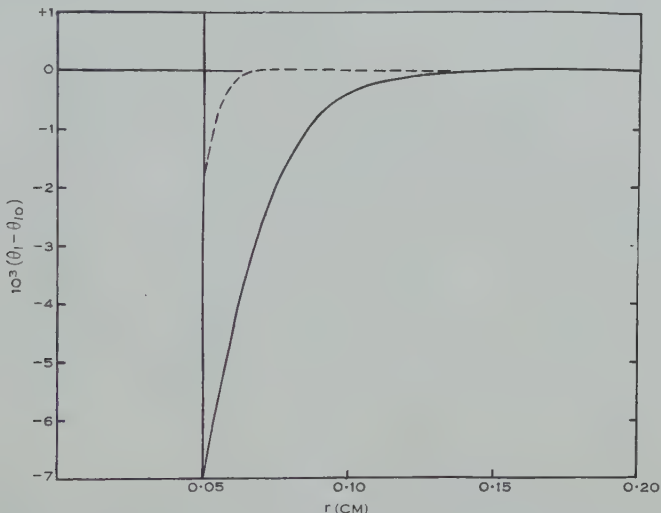


Fig. 2.—Theoretical values of change in moisture content, $\theta_l - \theta_{l_0}$, in relation to radial distance from probe axis, r . Curves for Yolo light clay at $20 \text{ }^\circ\text{C}$ with $\theta_{l_0} = 0.082$ and for $Q/4\pi\lambda = 0.1 \text{ }^\circ\text{C}$. Full curve for $t = 180$ sec, broken curve for $t = 10$ sec.

In the present example the calculated redistribution of moisture is almost negligible. Since the calculations were performed for the moisture content corresponding to the maximum of D_T/D_θ it could be expected that the moisture changes would be even less for other θ_l -values and also for the coarser-textured soil. This was confirmed by further calculations.

Larger values of $\theta_l - \theta_{l_0}$ are obtained for smaller R -values and for higher average temperatures. It follows from equations (27) and (29) that at $r = R$ the value of $\theta_l - \theta_{l_0}$ is approximately inversely proportional to R . Thus with a single heating wire of radius 0.005 cm (cf. Part I) the drying of the soil near the wire would be appreciable in the foregoing example. With moist soils it is therefore advisable not to reduce the probe diameter much below 0.1 cm.

The increase of $\theta_i - \theta_{i0}$ with increasing average temperature is principally due to the fact that both D_{T_v} and D_{θ_v} increase rapidly with increasing temperature and (over a limited temperature range) roughly in the same proportion. It follows further from equations (27) and (29) that $\theta_i - \theta_{i0}$ is approximately proportional to $D_T/D_0^{\frac{1}{2}}$ for $r=R$. In our numerical example, for instance, we find, for $t=180$ sec, at 60°C a value of -0.018 for $\theta_i - \theta_{i0}$ against -0.007 at 20°C .

With soils of finer texture than Yolo light clay somewhat larger values of $\theta_i - \theta_{i0}$ than those computed here must be expected. However, it seems improbable that even with very heavy clays the results will be greater by more than a factor five.

The effect of moisture movement on the temperature curve is also small, except at high average temperatures. The solution for $T - T_0$ is

$$T - T_0 = \frac{-b_1(u_2 - u_{2,0}) + b_2(u_1 - u_{1,0})}{-b_1 + b_2} = \frac{Q(1 + \varepsilon)}{4\pi\lambda'_2(1 - b_2/b_1)} [\ln(4\kappa_2 t/r^2) - \gamma + 0(\tau^{-1})], \quad \dots\dots\dots (30)$$

with

$$\varepsilon = -\frac{b_2(u_1 - u_{1,0})}{b_1(u_2 - u_{2,0})} > 0. \quad \dots\dots\dots (31)$$

Here λ'_2 differs little from λ and κ_2 little from a so that the temperature curve is little affected as long as $-b_2/b_1$ and ε are small in comparison with unity. In the above numerical example, for instance, we have at 20°C and $r=R=0.05$ cm: $-b_2/b_1 = 2.6 \times 10^{-5}$, $\varepsilon = 0.0027$ at $t=10$ sec and 0.0058 at $t=180$ sec; at 60°C : $-b_2/b_1 = 3.9 \times 10^{-4}$, $\varepsilon = 0.039$ at $t=10$ sec and 0.053 at $t=180$ sec. For the sand at 20°C with $\theta_i = 0.006$ (corresponding to the maximum of D_T/D_0 in Fig. 1) we find with $r=R=0.05$ cm: $-b_2/b_1 = 1.8 \times 10^{-3}$, $\varepsilon = 0.039$ at $t=10$ sec and 0.053 at $t=180$ sec.

The factor $1 + \varepsilon$ expresses the influence on the temperature rise of changes in heat transfer by vapour movement due to vapour flow under the influence of moisture gradients. It can be easily checked that as $t \rightarrow \infty$, $\lambda'_2(1 - b_2/b_1)/(1 + \varepsilon) \rightarrow \lambda - L\rho_l D_{\theta_v} D_T/D_0$. The latter value holds for steady-state heat conduction with $q_m = 0$ and with no influence of gravity (see de Vries 1958).

The results arrived at in this section are to be considered as approximate in view of the simplifying assumptions underlying the theory. The principal simplifications made are: (i) the exclusion of hysteresis, (ii) the assumption of constant coefficients, (iii) the neglect of gravity. Their influence will now be briefly discussed.

(i) *Exclusion of Hysteresis.*—With soils the effect of hysteresis will only be appreciable at relatively low moisture tensions where the changes in moisture content are very small. In the examples given above the effect of hysteresis will be negligible.

(ii) *Assumption of Constant Coefficients.*—It is impossible to assess the influence of the assumption of constant coefficients *a priori*. However, the fact that the approximate theory predicts only small variations in θ_i lends some

support to the procedure in retrospect. If we can consider the solution for constant coefficients as a reasonable first approximation to the exact solution, it can be shown by substitution that the neglected terms in equations (9) and (10), such as $(\nabla D_0 \nabla \theta_i)$, are small in comparison with other terms in these equations. We therefore believe that the theoretical values of $\theta_i - \theta_{i0}$ given here will at least be of the correct order of magnitude.

The situation is more complex when we consider the deviations of the temperature curve from that for the ideal case of non-varying λ and C . The drying of the soil near the probe will cause a decrease of these parameters, which is not taken into account in the present theory. This variation of the thermal properties will affect the temperature curve to a degree which may be greater than that expressed by the variation of ϵ in equation (30). For instance, in the example of Yolo light clay at 20 °C the calculated decrease of θ_i at the probe surface by an amount of 0.007 at 180 sec will cause a decrease there of 6 per cent. in λ and of 2 per cent. in C . These effects become rapidly smaller with increasing distance from the probe, but the total effect on the value of $T - T_0$ could well be of the same order as or greater than the influence of ϵ , which at 180 sec is 0.58 per cent. Both the variation of λ and of ϵ with time will cause a curvature in the $T - T_0$ against $\ln t$ relation that is convex toward the $\ln t$ axis.

(iii) *Neglect of Gravity*.—The influence of gravity is also only appreciable at low moisture tensions (see Section V); it is negligible in the examples given above.

Summarizing, we conclude that for most soils the redistribution of moisture and its effect on the temperature rise will be slight at average temperatures below about 40 °C for probe experiments of the type considered here, i.e. for a probe diameter of 0.1 cm, a heating time of the order of 100 sec, and a temperature rise at the probe surface of less than 1 °C.

V. THE INFLUENCE OF GRAVITY

Gravity affects the experiments in two ways. Firstly it causes uneven initial distribution of moisture in the soil and secondly it promotes moisture movement in a downward direction after the equilibrium is disturbed.

The condition for isothermal equilibrium is

$$\Phi = \Psi(\theta_i) + z = \Psi_0, \dots\dots\dots (32)$$

where Ψ_0 is the Ψ -value at $z=0$, e.g. at the probe centre. The moisture distribution follows from (32) through the relation between θ_i and Ψ .

In an experiment with a horizontal probe z varies practically from $-\frac{1}{2}R_1$ to $+\frac{1}{2}R_1$, when R_1 is the radius of the soil cylinder around the probe at which the heat flux is less than a small fraction (say 3 per cent.) of that at the probe. In Part I we have already shown that R_1 is about 5 cm in our experiments.

The variation of θ_i over a height of 10 cm will be negligible at low moisture contents, where $-\Psi \gg 10$ cm and $\partial \theta_i / \partial \Psi$ is small. With the Yolo light clay this is the case for $\theta_i < 0.35$ and with the sand for $\theta_i < 0.05$. At higher values of θ_i the moisture content at the height of the probe will be close to the average

moisture content in the sample when the relation between θ_i and Ψ can be considered as linear in the region Ψ_0-5 to Ψ_0+5 .

With such a non-radial initial moisture distribution the λ -value found from a probe experiment will usually correspond to a moisture content that is slightly higher than the average one, because the better conducting lower half of the sample will contribute more to the heat transfer than the poorer conducting upper half. However, it will be noted that at high moisture contents both λ and a vary slowly with θ_i .

The influence of gravity on the moisture movement, as expressed by the term $\partial K/\partial z$ in equation (9), is small over the entire moisture range. At low moisture contents this is due to the small values of K and $\partial K/\partial \theta_i$, at higher moisture contents to the small changes in θ_i .

Summarizing, we can say that the possible importance of the influence of gravity can be judged from a consideration of the variation of θ_i , λ , and a in a Ψ -region with a width of about 10 cm and with the Ψ -value at the probe axis (supposed to be horizontal) as its mid point. In practical cases this influence will often be small and the experimental λ -value will then apply to the average moisture content in the sample.

A quantitative treatment can only be given on the basis of an analytical or numerical solution of equations (9) and (10). We have not attempted such an approach.

VI. EXPERIMENTAL DATA

The theoretical conclusions of the previous section are generally confirmed by experimental experience. At sufficiently high moisture contents the theory for non-varying conductivity and diffusivity applies to probe experiments. However, one must be aware of complications due to an uneven moisture distribution under the influence of gravity.

In the field the water pressure in the upper soil layers is usually less than -100 cm and the influence of gravity then is negligible. Low values of $|\Psi|$ occur, for instance, during infiltration after rain or irrigation. In that case the last term on the right-hand side of equation (10), which represents transfer of sensible heat by liquid movement, is not negligible and the simple probe theory no longer applies. Very large apparent values of λ have been observed under conditions of infiltration.

Results of a laboratory experiment at 20°C on Leighton Buzzard sand (particle sizes 0.060 – 0.085 cm) with a dry density of 1.54 g cm $^{-3}$ and an average moisture content $\theta_i=0.22$ are shown in Figure 3. In the sample (height= 10 cm) the moisture content varied almost linearly from about 0.36 at the bottom to about 0.07 at the top. No apparent deviations from the behaviour of a homogeneous soil are observed. The experiment leads to $\lambda=5.5 \times 10^{-3}$ cal cm $^{-1}$ sec $^{-1}$ $^\circ\text{C}^{-1}$ from which we calculate $a=1.13 \times 10^{-2}$ cm 2 sec $^{-1}$. From equation (33) of Part I with $R=0.055$ cm we find $r=0.026$ cm. This is somewhat greater than the value of 0.021 found for the same probe in dry soil (see Part I). The difference is probably due to the uneven moisture distribution. Another experi-

ment on the same soil with a single heating wire gave the same value of λ within the experimental degree of accuracy.

Results for the same probe and sand now with $\theta_l = 0.023$, dry density 1.54 g cm^{-3} , and temperature 20°C are given in Figure 4. Here a slight curvature due to moisture movement can be noticed. The points for the cooling branch were found by using the extrapolated broken curve for the heating branch. The drying of the soil in the immediate vicinity of the probe results in a lower conductivity and therefore a more rapid fall in the first stages of cooling than with constant conductivity. The cooling branch lies therefore above the heating

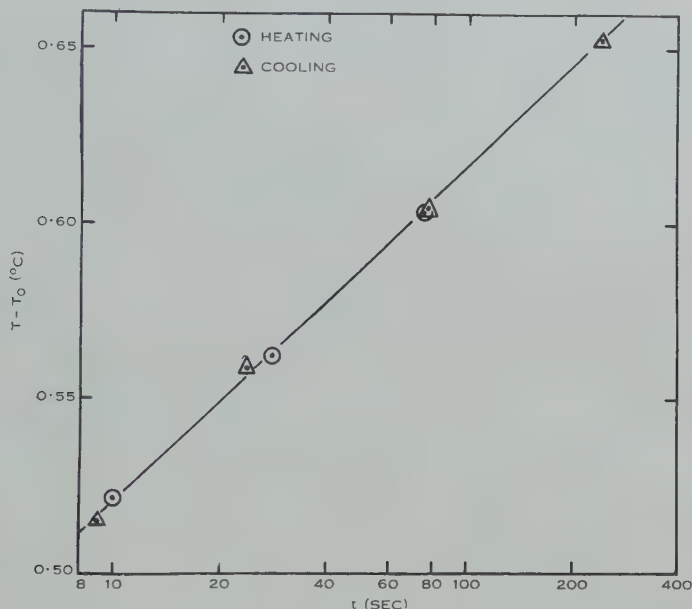


Fig. 3.—Temperature rise, $T - T_0$, against time for a probe experiment with a coarse sand at volumetric moisture content $\theta_l = 0.22$. Heat input $Q = 2.90 \times 10^{-3} \text{ cal cm}^{-1} \text{ sec}^{-1}$.

branch.* It must be noted that the scale of ordinates does not extend to zero and that the relative differences between the two branches are only of the order of a few per cent. The difference at 10 sec amounts to a distance of 0.8 mm on the galvanometer scale.

It is quite possible to draw a straight line through the points of the heating branch. If no cooling branch had been observed the effect of moisture movement would have remained unnoticed. From the straight line passing through the lower two points of the heating branch we find $\lambda = 2.8 \times 10^{-3} \text{ cal cm}^{-1} \text{ sec}^{-1} ^\circ \text{C}^{-1}$ and $r = 0.022 \text{ cm}$; from the straight line through the cooling points, $\lambda = 3.2 \times 10^{-3} \text{ cal cm}^{-1} \text{ sec}^{-1} ^\circ \text{C}^{-1}$ and $r = 0.018 \text{ cm}$. The former values are to be considered as the more reliable.

* The theory of Section IV cannot deal effectively with the cooling branch because of the assumption of constant λ and C .

It follows that in the region where moisture movement has a noticeable effect the λ -value obtained from a probe experiment will depend to a certain degree on the exact experimental conditions (e.g. the values of R and Q). When the probe method is used to measure the moisture content of soil it is therefore of importance to standardize the experimental procedure as far as possible, so that, for instance, the conditions are similar during calibration and actual measurements.

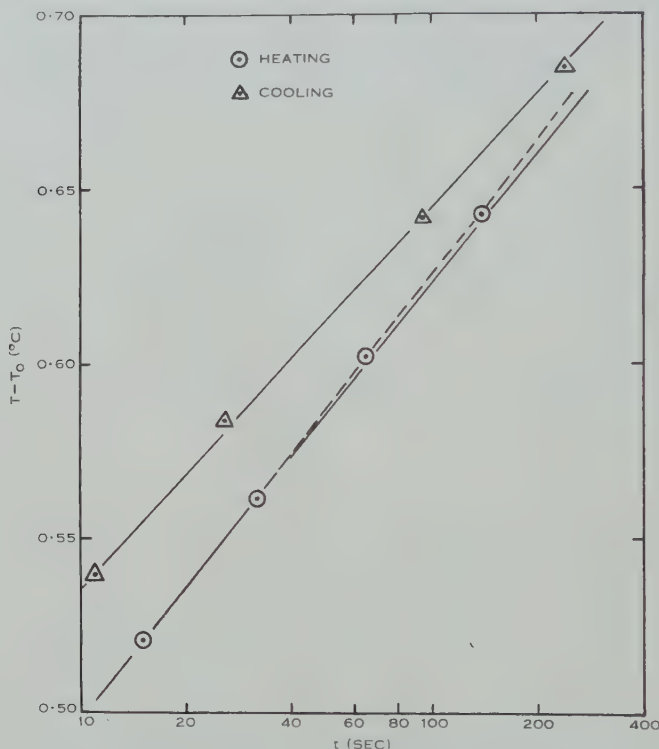


Fig. 4.—Temperature rise, $T - T_0$, against time for a probe experiment with a coarse sand at volumetric moisture content $\theta_t = 0.023$. Heat input $Q = 1.86 \times 10^{-3} \text{ cal cm}^{-1} \text{ sec}^{-1}$.

The effect of moisture movement could be minimized by feeding moisture into the soil through the probe wall at a required rate during heating. Experimental work along this line is in hand.

VII. CONCLUSIONS

In conclusion we can say that probe experiments of proper design will yield a reliable value of the thermal conductivity of moist soils. The attainable degree of accuracy will be sufficient for application in the field and for most laboratory investigations.

When it is intended to measure the soil moisture content by a probe method a high degree of accuracy is required and the complications discussed in this paper have to be taken into account.

The conclusion of Part I, that no high degree of accuracy can be expected in measuring the thermal diffusivity by the probe method, holds to an enlarged degree for moist soils.

It must further be kept in mind that heat transfer in a moist porous material not only depends on its thermal characteristics and on the thermal initial and boundary conditions, but also (be it often to a small extent) on the moisture characteristics and conditions for moisture flow. The simultaneous transfer of heat and moisture is governed by the differential equations given in Section III together with proper initial and boundary conditions.

VIII. REFERENCES

- BAVER, L. D. (1956).—"Soil Physics." 3rd Ed. p. 251. (John Wiley and Sons: New York.)
- CARSLAW, H. S., and JAEGER, J. C. (1950).—"Conduction of Heat in Solids." (Clarendon Press: Oxford.)
- CRANK, J. (1956).—"The Mathematics of Diffusion." p. 309. (Clarendon Press: Oxford.)
- PHILIP, J. R. (1955).—The concept of diffusion applied to soil water. *Proc. Nat. Acad. Sci. India*. **24A**: 93–104.
- PHILIP, J. R. (1957a).—The theory of infiltration, 2. The profile at infinity. *Soil Sci.* **83**: 435–48.
- PHILIP, J. R. (1957b).—Evaporation, and moisture and heat fields in the soil. *J. Met.* **14**: 354–66.
- PHILIP, J. R., and DE VRIES, D. A. (1957).—Moisture movement in porous materials under temperature gradients. *Trans. Amer. Geophys. Union* **38**: 222–32, 594.
- DE VRIES, D. A. (1958).—Simultaneous transfer of heat and moisture in porous media. *Trans. Amer. Geophys. Union* **39**. (In press.)
- DE VRIES, D. A., and PECK, A. J. (1958).—On the cylindrical probe method of measuring thermal conductivity with special reference to soils. I. Extension of theory and discussion of probe characteristics. *Aust. J. Phys.* **11**: 255–71.

ELECTROMAGNETIC RADIATION FROM ELECTRONS ROTATING IN AN IONIZED MEDIUM UNDER THE ACTION OF A UNIFORM MAGNETIC FIELD

By R. Q. TWISS* and J. A. ROBERTS*

[Manuscript received February 27, 1958]

Summary

A theory is given for the radiation from a fast electron rotating, under the action of an external magnetic field, in an ionized plasma. It is shown that, although the radiation is emitted predominantly in the extraordinary mode, the ordinary mode is also weakly excited, even in the limiting case in which the density of the background plasma is vanishingly small. At the *harmonics* of the gyro frequency of the fast electron the power radiated in the ordinary mode is a few per cent. of that radiated in the extraordinary mode. This ratio is independent of v_0 , the velocity of the fast electron, as long as v_0 is sufficiently small compared with c , the velocity of light. However, at the *fundamental* gyro frequency the power radiated in the ordinary mode is lower by a factor $\simeq 10^{-2}(v_0/c)^4$ than that radiated in the extraordinary mode and indeed is significantly smaller than that radiated, in either mode, at the third harmonic.

The gyro theory of the non-thermal radiation from the Sun is discussed in the light of these results and it is argued that this mechanism cannot explain the phenomena associated with the bursts of spectral types II and III. However, it is conceivable that the radiation on spectral type I may be of gyro origin, though even in this case there are serious objections to this explanation.

I. INTRODUCTION

The possibility that the non-thermal radiation from the Sun might be due to gyro radiation from fast electrons was suggested at a very early stage in the development of radio astronomy (Kiepenheuer 1946), but in its initial form this hypothesis seemed subject to several very serious objections. The most important of these was that gyrating electrons would radiate only in the so-called extraordinary mode. At the fundamental frequency at which, it was thought, the majority of the radio energy would be generated, it was argued that this mode could not escape from the Sun (Ryle 1950). It has recently been suggested (Kruse, Marshall, and Platt 1956) that this objection is not conclusive, since in fact a gyrating electron will radiate both in the ordinary and in the extraordinary mode except in directions parallel and perpendicular to the magnetic field. However, in estimating the magnitude of the radiation that might be expected in a practical case, Kruse, Marshall, and Platt assumed that an electron gyrating in a plasma would radiate as much energy as in free space, and that all the energy would be in the ordinary mode. These assumptions are not valid, and before one

* Division of Radiophysics, C.S.I.R.O., University Grounds, Chippendale, N.S.W.

can decide whether gyro radiation plays any significant part in the non-thermal solar radio emission one must have a quantitative theory for radiation by electrons in an ionized medium.

In the present paper we consider the idealized case of a single charged particle of arbitrary energy rotating in a circular orbit in a uniform ionized medium of zero temperature and we shall assume that the motions of the ions are too small to be significant. The more general case where the ionized medium is at a finite temperature and where there is a distribution of fast rotating particles with arbitrary axial velocities has yet to be analysed.

II. THE FUNDAMENTAL EQUATIONS AND METHOD OF SOLUTION

In the small signal theory in which the quadratic terms involving the time-dependent field and space-charge quantities are negligible, the Maxwell-Lorentz equations which determine the generation and propagation of electromagnetic waves in an ionized medium become linear. When the ionized medium is a neutral plasma at zero temperature with infinitely massive ions acted upon by a uniform external magnetic field, then the fundamental equations determining the radiation from a rotating charged particle consist of the vector Maxwell equations

$$\nabla \wedge \mathbf{E} = -\partial \mathbf{B} / \partial t; \quad \nabla \wedge \mathbf{H} = \rho_0 \mathbf{v} + \partial \mathbf{D} / \partial t + \mathbf{I}_0, \quad \dots \dots \dots (1)$$

where

$$\mathbf{B} = \mu_0 \mathbf{H}, \quad \mathbf{D} = \epsilon_0 \mathbf{E}, \quad \dots \dots \dots (2)$$

the equation representing the conservation of charge

$$\partial \rho / \partial t + \rho_0 \nabla \cdot \mathbf{v} = 0, \quad \dots \dots \dots (3)$$

and the Lorentz force equation

$$\partial(m_0 \mathbf{v}) / \partial t = -e(\mathbf{E} + \mathbf{v} \wedge \mathbf{a} B_0). \quad \dots \dots \dots (4)$$

Here

ρ_0, ρ are the average and time-dependent charge densities respectively of the background plasma,

\mathbf{v} is the time-dependent velocity of the background plasma electrons,

$-e, m_0$ are the electronic charge and rest mass respectively,

B_0 is the flux density of the external axial magnetic field,

\mathbf{a} is the unit axial vector $(0, 0, 1)$, and

\mathbf{I}_0 is the current density associated with the rotating charged particle.

These equations are subject to the boundary condition that the solution should consist only of outward going waves at distances sufficiently far from the origin.

The solution which we shall give here is based directly upon the analysis given in an earlier paper (Twiss 1952), hereafter referred to as I, in which it was shown that the general solutions of the Maxwell-Lorentz equations in an ionized

medium composed of arbitrarily many beams of charged particles and pervaded by a uniform external magnetic field could be represented by an integral expansion over a complete orthogonal set of elementary vectors. These vectors are given by the orthogonal triad \mathbf{L} , \mathbf{M} , \mathbf{N} , where

$$\mathbf{L} = \nabla \psi(x^1, x^2), \quad \mathbf{M} = \nabla \psi(x^1, x^2) \wedge \mathbf{a}, \quad \mathbf{N} = \psi(x^1, x^2) \mathbf{a}, \quad \dots \quad (5)$$

where the scalar quantity $\psi(x^1, x^2)$ satisfies the two-dimensional wave equation

$$\nabla^2 \psi(x^1, x^2) + p^2 \psi(x^1, x^2) = 0. \quad \dots \dots \dots (6)$$

Here (x^1, x^2) are the coordinates transverse to the magnetic field and p , the transverse wave number, is a real quantity.

In the case of radiation by a rotating charged particle it is natural to take the origin at the centre of rotation and use circular cylindrical coordinates (r, φ, z) with the plane $z=0$ as the plane of rotation. Then, since the observable quantities will be finite on the axis $r=0$ we may, without loss of generality, take

$$\psi(x^1, x^2) \equiv \psi_{p,m}(r, \varphi) = J_m(pr) \exp(-im\varphi) \quad \dots \dots \dots (7)$$

as a typical solution of equation (6).

Subject to certain conditions of boundedness and integrability, which will be valid in any physically realizable case, an arbitrary vector field $\mathbf{F}(r, \varphi, z, t)$ can be expanded by

$$\begin{aligned} \mathbf{F}(r, \varphi, z, t) = & \sum_{m=-\infty}^{\infty} \int_0^{\infty} dp [F_1(p, m, z, t) \nabla \psi_{p,m} \\ & + F_2(p, m, z, t) \nabla \psi_{p,m} \wedge \mathbf{a} + F_3(p, m, z, t) \psi_{p,m} \mathbf{a}], \quad \dots \quad (8) \end{aligned}$$

where $\psi_{p,m}(r, \varphi)$ is given by equation (7). Similarly, an arbitrary scalar field $U(r, \varphi, z, t)$ can be expanded in the form

$$U(r, \varphi, z, t) = \sum_{m=-\infty}^{\infty} \int_0^{\infty} dp u_0(p, m, z, t) \psi_{p,m}. \quad \dots \quad (9)$$

The quantities $F_l(p, m, z, t)$ ($l=1, 2, 3$) and $u_0(p, m, z, t)$ can be found explicitly in terms of $\mathbf{F}(r, \varphi, z, t)$ and $U(r, \varphi, z, t)$ by the Fourier-Bessel inversion formulae which state that if

$$G(p) = \int_0^{\infty} r g(r) J_m(pr) dr, \quad \dots \dots \dots (10)$$

then

$$g(r) = \int_0^{\infty} p G(p) J_m(pr) dp, \quad \dots \dots \dots (11)$$

and vice versa, provided that

$$\int_0^{\infty} r^{\frac{1}{2}} |g(r)| dr$$

exists in the sense of Lebesgue.

The method of solution given in I is to express the vector quantities \mathbf{E} , \mathbf{H} , \mathbf{v} , and \mathbf{I}_0 which appear in equations (1)-(4) by the vector expansions

$$E(r, \varphi, z, t) = \sum_{m=-\infty}^{\infty} \int_0^{\infty} dp [E_1(p, m, z, t) \nabla \psi_{p,m} + E_2(p, m, z, t) \nabla \psi_{p,m} \wedge \mathbf{a} + E_3(p, m, z, t) \psi_{p,m} \mathbf{a}], \quad \dots \quad (12a)$$

$$H(r, \varphi, z, t) = \sum_{m=-\infty}^{\infty} \int_0^{\infty} dp [H_1(p, m, z, t) \nabla \psi_{p,m} + H_2(p, m, z, t) \nabla \psi_{p,m} \wedge \mathbf{a} + H_3(p, m, z, t) \psi_{p,m} \mathbf{a}], \quad \dots \quad (12b)$$

$$v(r, \varphi, z, t) = \sum_{m=-\infty}^{\infty} \int_0^{\infty} dp [v_1(p, m, z, t) \nabla \psi_{p,m} + v_2(p, m, z, t) \nabla \psi_{p,m} \wedge \mathbf{a} + v_3(p, m, z, t) \psi_{p,m} \mathbf{a}], \quad \dots \quad (12c)$$

$$I_0(r, \varphi, z, t) = \sum_{m=-\infty}^{\infty} \int_0^{\infty} dp [I_1(p, m, z, t) \nabla \psi_{p,m} + I_2(p, m, z, t) \nabla \psi_{p,m} \wedge \mathbf{a} + I_3(p, m, z, t) \psi_{p,m} \mathbf{a}], \quad \dots \quad (12d)$$

and the scalar quantity $\rho(r, \varphi, z, t)$ by the scalar expansion

$$\rho(r, \varphi, z, t) = \sum_{m=-\infty}^{\infty} \int_0^{\infty} dp \sigma(p, m, z, t) \psi_{p,m}. \quad \dots \quad (13)$$

Now because of equation (6) it was shown in I that all the vector-yielding operations in equations (1)-(4) leave the general form of a vector

$$\mathbf{V} = a(z, t) \mathbf{L} + b(z, t) \mathbf{M} + c(z, t) \mathbf{N}$$

invariant, while all the scalar-yielding operations in equations (1)-(4) when applied to a vector \mathbf{V} reduce to a form linearly dependent upon $\psi_{p,m}$. Hence all the vector equations of equations (1)-(4) can be expressed in the general form

$$\sum_{m=-\infty}^{\infty} \int_0^{\infty} dp [\alpha(p, m, z, t) \mathbf{L} + \beta(p, m, z, t) \mathbf{M} + \gamma(p, m, z, t) \mathbf{N}] = 0,$$

while all the scalar equations are of the form

$$\sum_{m=-\infty}^{\infty} \int_0^{\infty} dp \omega(p, m, z, t) \psi_{p,m} = 0.$$

If these equations are to be satisfied for all z, t we must have

$$\alpha(p, m, z, t) = \beta(p, m, z, t) = \gamma(p, m, z, t) = \omega(p, m, z, t) = 0, \quad \dots \quad (14)$$

and these latter equations can be solved for the components $E_l(p, m, z, t)$ ($l=1, 2, 3$) of the electric field in terms of the components of the "driving function" $I_0(r, \varphi, z, t)$ and of the initial conditions along the lines given in I.

III. THE VECTOR EXPANSION FOR THE DRIVING FUNCTION

In terms of the Dirac δ -function we may write the "driving function" $I_0(r, \varphi, z, t)$ in the form

$$I_0(r, \varphi, z, t) = - \sum_{n=0}^{\infty} e_s(z) \delta(r - r_0) \delta(t - nt_0 - t_0 \varphi / 2\pi) \mathbf{b}, \quad \dots \quad (15)$$

where

t_0 is the rotation period defined by

$$2\pi/t_0 = |e_s B_0/m_s| = |\mathbf{v}_0|/r_0, \quad \dots\dots\dots (16)$$

\mathbf{v}_0 is the velocity of the particle, with charge $-e_s$, in its orbit of radius r_0 ,

\mathbf{b} is the unit vector $(0, 1, 0)$, and

m_s is the relativistic transverse mass of the rotating particle.

The summation in equation (15) arises because φ is confined to the range $0 \leq \varphi < 2\pi$ by the need to keep all functions of position single valued; we have assumed that the fast particle is "started up" at time $t=0$ so that at a later stage we can use the Laplace transform analysis.

In a practical case e_s, m_s will have the values appropriate to the electron, but we shall not make this restriction to begin with.

As a first step in the solution outlined in the previous section, we must find explicit expressions for $I_l(p, m, z, t)$ ($l=1, 2, 3$) in equation (12).

From equation (15), we have immediately

$$I_3(p, m, z, t) = 0.$$

To find I_1 , we operate on both sides of equation (12) with the scalar operator

$$(\nabla - \mathbf{a}\partial/\partial z),$$

and get

$$(\nabla - \mathbf{a}\partial/\partial z) \cdot \mathbf{I}_0(r, \varphi, z, t) = \sum_{m=-\infty}^{\infty} \int_0^{\infty} dp I_1(p, m, z, t) \nabla^2 \psi_{p,m}.$$

Now from equations (6), (10), (11), and the Fourier series inversion formulae, we have immediately

$$\begin{aligned} I_1(p, m, z, t) &= \frac{-1}{2\pi p} \int_0^{2\pi} d\varphi \int_0^{\infty} dr \, r \left[\frac{1}{r} \frac{\partial}{\partial \varphi} \left\{ - \sum_{n=0}^{\infty} e_s \delta(z) \delta(r-r_0) \right. \right. \\ &\quad \left. \left. \times \delta\left(t - nt_0 - \frac{t_0 \varphi}{2\pi}\right) \right\} J_m(pr) \exp(im\varphi) \right] \\ &= - \frac{e_s \delta(z) J_m(pr_0)}{2\pi p} \cdot \frac{2\pi im}{t_0} \exp\left(\frac{2\pi imt}{t_0}\right), \quad t \geq 0. \quad \dots (17) \end{aligned}$$

To find $I_2(p, m, z, t)$ we take the curl of both sides of equation (12d) and then take the scalar product with \mathbf{a} to get

$$\mathbf{a} \cdot [\nabla \wedge \mathbf{I}_0(r, \varphi, z, t)] = - \sum_{m=-\infty}^{\infty} \int_0^{\infty} dp I_2(p, m, z, t) \nabla^2 \psi_{p,m},$$

whence

$$\begin{aligned} I_2(p, m, z, t) &= \frac{1}{2\pi p} \int_0^{2\pi} d\varphi \int_0^{\infty} dr \, r \left[\frac{1}{r} \frac{\partial}{\partial r} \left\{ -r \sum_{n=0}^{\infty} e_s \delta(z) \delta(r-r_0) \right. \right. \\ &\quad \left. \left. \times \delta\left(t - nt_0 - \frac{t_0 \varphi}{2\pi}\right) \right\} J_m(pr) \exp(im\varphi) \right] \\ &= - \frac{e_s \delta(z) p r_0 J'_m(pr_0)}{2\pi p} \cdot \frac{2\pi}{t_0} \exp\left(\frac{2\pi imt}{t_0}\right), \quad t \geq 0. \quad \dots (18) \end{aligned}$$

IV. THE SOLUTION IN TERMS OF THE DRIVING FUNCTION

If we substitute from equations (12), (13) in equations (1)-(4), then from equations of the form of equation (14) it can be shown that the Maxwell equations reduce to

$$\left. \begin{aligned} -\mu_0 \partial H_1 / \partial t &= \partial E_2 / \partial z, & \partial H_2 / \partial z &= \rho_0 v_1 + \varepsilon_0 \partial E_1 / \partial t + I_1, \\ -\mu_0 \partial H_2 / \partial t &= E_3 - \partial E_1 / \partial z, & -\partial H_1 / \partial z + H_3 &= \rho_0 v_2 + \varepsilon_0 \partial E_2 / \partial t + I_2, \\ -\mu_0 \partial H_3 / \partial t &= p^2 E_2, & p^2 H_2 &= \rho_0 v_3 + \varepsilon_0 \partial E_3 / \partial t. \end{aligned} \right\} \dots (19)$$

From the charge conservation equation we get

$$\partial \sigma / \partial t = -\rho_0 (-p^2 v_1 + \partial v_3 / \partial z), \dots (20)$$

and from the Lorentz force equations

$$\left. \begin{aligned} \partial(m_0 v_1) / \partial t &= -e E_1 + m_0 \omega_H v_2, \\ \partial(m_0 v_2) / \partial t &= -e E_2 - m_0 \omega_H v_1, \\ \partial(m_0 v_3) / \partial t &= -e E_3, \end{aligned} \right\} \dots (21)$$

where

$$\omega_H = e B_0 / m_0 \dots (22)$$

is the gyro angular frequency of the background plasma electrons.

Except for the forcing terms I_1, I_2 in equation (19) this set of equations is a special case of those derived in I, in which the solution was obtained in terms of arbitrary initial conditions by means of successive Laplace transforms firstly with respect to the time coordinate t and secondly with respect to the axial coordinate z . To apply this procedure to the present case of an unbounded medium we introduce a cut over the plane $z=0$ and consider the solutions in the regions $z>0, z<0$ separately.

The double Laplace transform $F^{z,t}(k, \omega)$ of a quantity $F(z, t)$ can be defined in the region $z>0, t>0$ by

$$F^{z,t}(k, \omega) = \int_0^\infty dz \exp(-ikz) \int_0^\infty dt F(z, t) \exp(-i\omega t), \dots (23)$$

where

$$F(z, t) = \frac{1}{(2\pi)^2} \int_{-\infty - i\gamma_1}^{\infty - i\gamma_1} dk \exp(ikz) \int_{-\infty - i\gamma_2}^{\infty - i\gamma_2} d\omega F^{z,t}(k, \omega) \exp(i\omega t), \dots (24)$$

and where γ_1, γ_2 are real positive numbers such that all the poles and singularities of $F^{z,t}(k, \omega)$ and $F^t(k, \omega)$ lie above the lines

$$\text{Im}(k) + \gamma_1 = 0, \quad \text{Im}(\omega) + \gamma_2 = 0$$

in the complex k - and ω -planes respectively.

In the region $z < 0$, $t > 0$, $F^{z,t}(k, \omega)$ is defined by a similar equation in which all the singularities of $F^{z,t}(k, \omega)$ lie below the line

$$\text{Im}(k) + \gamma'_1 = 0$$

in the complex k -plane.

Now that the plane $z=0$ has been excluded by a cut, the quantities I_1, I_2 in equation (19), may be taken as zero in the regions $z > 0$ and $z < 0$ and we can use the solution, derived in I, for $E_l^{z,t}(p, m, k, \omega)$ ($l=1, 2, 3$) in terms of the initial conditions. In the present case the disturbance at $t=0$ may be taken as identically zero in the regions $z > 0$ and $z < 0$ and under these conditions we have immediately from equation (3.52) of I

$$E_l^{z,t}(k, \omega) = \sum_{r=1}^3 C_r^*(p, m, k, \omega) \mathbf{A}_{lr}(p, m, k, \omega) / \det \mathbf{A}. \quad \dots (25)$$

Here $\mathbf{A}_{lr}(p, m, k, \omega)$ is the minor of $a_{lr}(p, m, k, \omega)$ in the matrix \mathbf{A} defined by equation (3.47) in I which, in the present case, may be written

$$\mathbf{A} = \begin{bmatrix} k^2 - \frac{\omega^2}{c^2} + \frac{\omega^2 \omega_0^2}{c^2(\omega^2 - \omega_H^2)}, & \frac{-i\omega \omega_H \omega_0^2}{c^2(\omega^2 - \omega_H^2)}, & ik \\ \frac{i\omega \omega_H \omega_0^2}{c^2(\omega^2 - \omega_H^2)}, & p^2 + k^2 - \frac{\omega^2}{c^2} + \frac{\omega^2 \omega_0^2}{c^2(\omega^2 - \omega_H^2)}, & 0 \\ ikp^2, & 0, & \frac{\omega^2 - \omega_0^2}{c^2} - p^2 \end{bmatrix}, \quad \dots (26)$$

where

$$\omega_0 = (-e\rho_0/\epsilon_0 m_0)^{\frac{1}{2}} \quad \dots (27)$$

is the angular plasma frequency associated with the background electrons.

The column matrix $C^*(p, m, k, \omega)$ defined by equation (3.49) of I reduces in the present case to the simple form given, for $z > 0$, by

$$\left. \begin{aligned} C_1^*(p, m, k, \omega) &= -ikE_1^t(+0) - i\omega\mu_0 H_2^t(+0), \\ C_2^*(p, m, k, \omega) &= -ikE_2^t(+0) + i\omega\mu_0 H_1^t(+0), \\ C_3^*(p, m, k, \omega) &= p^2 E_1^t(+0). \end{aligned} \right\} \quad \dots (28)$$

It therefore depends only upon the transverse components of the electric and magnetic fields at $z=+0$.

The expression for $C^*(p, m, k, \omega)$ in the region $z < 0$ is of exactly similar form and involves the initial conditions $E_1^t(-0)$, $E_2^t(-0)$, $H_1^t(-0)$, and $H_2^t(-0)$.

The transverse fields in the region $z > 0$ and $z < 0$ are connected by the continuity conditions which, for the electric field components, give

$$\left. \begin{aligned} E_1^t(+0) &= E_1^t(-0), \\ E_2^t(+0) &= E_2^t(-0). \end{aligned} \right\} \quad \dots (29)$$

The continuity conditions on the tangential magnetic fields are complicated by the fact that the driving function consists of a current sheet in the plane $z=0$. However, by applying Stokes's theorem

$$\oint \mathbf{H} \cdot d\mathbf{s} = \int \mathbf{J} \cdot d\mathbf{S}.$$

it can be shown that

$$\left. \begin{aligned} H_1^t(+0) - H_1^t(-0) &= + \int_{-0}^{+0} dz I_2^t(p, m, z, \omega) \\ &= \frac{-e_s}{2\pi p i} \cdot \frac{2\pi/t_0}{\omega - 2\pi m/t_0} p r_0 J'_m(p r_0), \\ H_2^t(+0) - H_2^t(-0) &= - \int_{-0}^{+0} dz I_1^t(p, m, z, \omega) \\ &= \frac{+e_s}{2\pi p} \cdot \frac{2\pi m/t_0}{\omega - 2\pi m/t_0} J_m(p r_0). \end{aligned} \right\} \dots (30)$$

Two further relations between $E_1^t(+0)$, $E_2^t(+0)$, $H_1^t(+0)$, and $H_2^t(+0)$ are provided by the boundary conditions which require that the solution in the region $z > 0$ should consist only of outward-going waves. Two similar relations will exist between the transverse field components at $z = -0$ but in the present case these need not be set up explicitly since, from symmetry considerations, we have at once

$$\left. \begin{aligned} H_1^t(+0) &= -H_1^t(-0) = -\frac{1}{2} \cdot \frac{e_s}{2\pi p i} \cdot \frac{2\pi/t_0}{\omega - 2\pi m/t_0} p r_0 J'_m(p r_0), \\ H_2^t(+0) &= -H_2^t(-0) = +\frac{1}{2} \cdot \frac{e_s}{2\pi p} \cdot \frac{2\pi m/t_0}{\omega - 2\pi m/t_0} J_m(p r_0). \end{aligned} \right\} \dots (31)$$

The components at $z = -0$ can therefore be eliminated by means of equations (29) and (31).

(a) *The Inversion from the Complex k -Plane onto the Real z -Axis*

In the complex k -plane the singularities of $E_i^t(p, m, k, \omega)$ consist of simple poles at the zero of $\det \mathbf{A}$. In the present case $\det \mathbf{A}$ may be expanded in the biquadratic form given by

$\det \mathbf{A} \equiv$

$$\begin{aligned} & k^4 \frac{\omega^2 - \omega_0^2}{c^2} + k^2 \left[\frac{\omega^2 - \omega_0^2}{c^2} \left\{ p^2 - \frac{2\omega^2}{c^2} \left(1 - \frac{\omega_0^2}{\omega^2 - \omega_H^2} \right) \right\} + \frac{p^2 \omega^2}{c^2} \left(1 - \frac{\omega_0^2}{\omega^2 - \omega_H^2} \right) \right] \\ & - \frac{\omega^2}{c^2} \left(\frac{\omega^2 - \omega_0^2}{c^2} - p^2 \right) \left\{ p^2 \left(1 - \frac{\omega_0^2}{\omega^2 - \omega_H^2} \right) - \frac{\omega^2}{c^2} \left(1 - \frac{\omega_0^2}{\omega^2 - \omega_H^2} \right)^2 + \frac{\omega_H^2}{c^2} \cdot \frac{\omega_0^4}{(\omega^2 - \omega_H^2)^2} \right\}, \\ & \dots \dots \dots (32) \end{aligned}$$

which may be written in the equivalent form

$$\det \mathbf{A} \equiv \{(\omega^2 - \omega_0^2)/c^2\}[(k - k_1)(k - k_2)(k - k_3)(k - k_4)]. \dots (33)$$

The inversion from the complex k -plane onto the real z -axis expresses $E_l^t(p, m, z, \omega)$ as the sum of the residues of

$$iE_l^{z,t}(p, m, k, \omega) \exp ikz$$

at the poles given by the zeros of $\det \mathbf{A}$. Hence $E_l^t(p, m, z, \omega)$ may be represented as the sum of four partial waves of the form

$$\sum_{r=-2}^2 G_r(k_r) \exp(ik_r z).$$

Two of these partial waves will be reverse waves excited only by reflection and their propagation constants will be designated by $+k_1, +k_2$, where for the reasons given in I, the sign of $k_1(\omega), k_2(\omega)$ must be chosen so that

$$\lim_{\omega \rightarrow \infty} \frac{k_1(\omega)}{\omega} = \lim_{\omega \rightarrow \infty} \frac{k_2(\omega)}{\omega} = +\frac{1}{c},$$

when ω tends to infinity along a line in the complex ω -plane which lies below all the singularities of $k_1(\omega), k_2(\omega)$.

To meet the boundary conditions at infinity the amplitudes $G_l(k_1), G_l(k_2)$ of these reverse partial waves must be identically zero. From equation (25) this implies that

$$\sum_{r=1}^3 C_r^*(p, m, k_s, \omega) \mathbf{A}_{lr}(p, m, k_s, \omega) = 0 \quad \dots \dots \dots (34)$$

for any† choice of l .

The amplitudes of the forward partial waves with propagation constants $-k_1, -k_2$ are proportional to

$$\sum_{r=1}^3 C_r^*(p, m, -k_s, \omega) \mathbf{A}_{lr}(p, m, -k_s, \omega), \quad s=1, 2.$$

By inspection of equation (26) we see that $\mathbf{A}_{l1}, \mathbf{A}_{l2}$ are even functions of k and \mathbf{A}_{l3} is an odd function of k for $l=1, 2$ and vice versa for $l=3$. Hence, from equations (28) and (34), we find that the amplitudes of the forward waves are proportional to

$$-2i\omega\mu_0[H_2^t(+0)\mathbf{A}_{l1}(p, m, -k_s, \omega) - H_1^t(+0)\mathbf{A}_{l2}(p, m, -k_s, \omega)], \quad s=1, 2, \dots \dots \dots (35)$$

† It is a well-known consequence of the theory of determinants that the three linear equations obtained by taking the three possible choices of l are linearly dependent when $\det \mathbf{A}(p, m, k_s, \omega) = 0$ so that this condition gives only two independent equations for $s=1, 2$.

and using equations (25), (26), (28), and (31) we get

$$\begin{aligned}
 E_1^t(p, m, z, \omega) = & - \sum_{s=1}^2 \frac{(\omega^2 - \omega_0^2)/c^2 - p^2}{(\omega^2 - \omega_0^2)/c^2} \cdot \frac{\omega \mu_0 e_s p m}{t_0 p^2 (\omega - 2\pi m/t_0)} \\
 & \times \left[\left\{ \left(p^2 + k_s^2 - \frac{\omega^2}{c^2} + \frac{\omega^2 \omega_0^2}{c^2 (\omega^2 - \omega_H^2)} \right) J_m(pr_0) \right. \right. \\
 & \left. \left. - \frac{\omega \omega_H \omega_0^2}{c^2 (\omega^2 - \omega_H^2)} \cdot \frac{pr_0 J'_m(pr_0)}{m} \right\} \frac{(-)^{s-1} \exp(-ik_s z)}{2k_s (k_1^2 - k_2^2)} \right], \\
 E_2^t(p, m, z, \omega) = & \sum_{s=1}^2 \frac{(\omega^2 - \omega_0^2)/c^2 - p^2}{(\omega^2 - \omega_0^2)/c^2} \cdot \frac{i \omega \mu_0 e_s p m}{t_0 p^2 (\omega - 2\pi m/t_0)} \\
 & \times \left[\left\{ \left(\frac{p^2 k_s^2}{(\omega^2 - \omega_0^2)/c^2 - p^2} + k_s^2 - \frac{\omega^2}{c^2} + \frac{\omega^2 \omega_0^2}{c^2 (\omega^2 - \omega_H^2)} \right) \frac{pr_0 J'_m(pr_0)}{m} \right. \right. \\
 & \left. \left. - \frac{\omega \omega_H \omega_0^2}{c^2 (\omega^2 - \omega_H^2)} J_m(pr_0) \right\} \frac{(-)^{s-1} \exp(-ik_s z)}{2k_s (k_1^2 - k_2^2)} \right], \\
 E_3^t(p, m, z, \omega) = & - \sum_{s=1}^2 \frac{p^2}{(\omega^2 - \omega_0^2)/c^2} \cdot \frac{i \omega \mu_0 e_s p m}{t_0 p^2 (\omega - 2\pi m/t_0)} \\
 & \times \left[\left\{ \left(p^2 + k_s^2 - \frac{\omega^2}{c^2} + \frac{\omega^2 \omega_0^2}{c^2 (\omega^2 - \omega_H^2)} \right) J_m(pr_0) \right. \right. \\
 & \left. \left. - \frac{\omega \omega_H \omega_0^2}{c^2 (\omega^2 - \omega_H^2)} \cdot \frac{pr_0 J'_m(pr_0)}{m} \right\} \frac{(-)^{s-1} \exp(-ik_s z)}{2k_s (k_1^2 - k_2^2)} \right].
 \end{aligned}
 \tag{36}$$

In the complex ω -plane the electric field components $E_l^t(p, m, z, \omega)$ ($l=1, 2, 3$) all have a simple pole at

$$\omega = 2\pi m/t_0 \equiv \omega_m, \quad \dots \tag{37}$$

that is, at the m th harmonic of the gyro frequency of the fast charged particle. The residue of this pole represents the steady state solution in which we are interested* in this paper. Thus, ignoring the decaying transient terms, we have

$$E_l(p, m, z, t) = i(\omega - 2\pi m/t_0) E_l^t(p, m, z, \omega) \exp(2\pi i m t/t_0). \quad \dots \tag{38}$$

V. THE ASYMPTOTIC SOLUTION AT LARGE DISTANCES FROM THE SOURCE

We shall consider the solution at very large distances from the source in a direction making an angle χ with the z -axis at the point

$$r = R \sin \chi, \quad z = R \cos \chi, \quad R \rightarrow \infty.$$

* It might appear from equation (36) that $E_l^t(p, m, z, \omega)$ would have poles at $\omega = \omega_0$ and $\omega = \omega_H$, that is, at the natural resonances associated with the plasma and gyro frequencies of the medium. However, a more detailed discussion involving the behaviour of $k_1(\omega)$, $k_2(\omega)$ at these frequencies shows that this is not the case. This latter result is to be expected on physical grounds since, if it did not hold, it would imply that the medium would oscillate with finite amplitude at its natural resonances whatever the nature of the initial stimulus. The transient terms, arising from the integrals around the cuts in the complex ω -plane that need to be inserted to eliminate the branch points of $E_l^t(p, m, z, \omega)$ decay with time and do not contribute to the steady state solution.

Since the solution has been derived for the case $z > 0$ we have

$$0 \leq \chi \leq \frac{1}{2}\pi, \quad 0 \leq \tan \chi \leq \infty.$$

Now as $r \rightarrow \infty$

$$\psi_{p,m}(r, \varphi) = J_m(pr) \exp(-im\varphi) \sim \{1/\sqrt{(\frac{1}{2}\pi pr)}\} \cos(pr - \frac{1}{2}m\pi - \frac{1}{4}\pi) \exp(-im\varphi), \quad \dots (39a)$$

as long as $p > 0$, while

$$\left. \begin{aligned} \nabla \psi_{p,m}(r, \varphi) &\sim (-ip\psi, 0, 0), \\ \nabla \psi_{p,m}(r, \varphi) \wedge \mathbf{a} &\sim (0, ip\psi, 0), \end{aligned} \right\} \dots \dots \dots (39b)$$

Since $(1/r)\partial\psi/\partial\varphi \sim r^{-3/2}$ and may be neglected.

Accordingly, the l th component of the electric field associated with the s th mode is of the general form

$$\begin{aligned} E_{ls}(R \sin \chi, m, R \cos \chi, t) \\ \sim \int_0^\infty \frac{dp W_{ls}(p)}{\sqrt{(\frac{1}{2}\pi p R \sin \chi)}} \cos(pR \sin \chi - \frac{1}{2}m\pi - \frac{1}{4}\pi) \\ \times \exp[-iR \cos \chi k_s(p)], \quad \dots \dots \dots (40) \end{aligned}$$

where s may take the value 1, 2 corresponding to the ordinary and extraordinary mode respectively, and where

$$(W_{1s}(p), W_{2s}(p), W_{3s}(p)) \equiv (-ipE_{1s}, ipE_{2s}, E_{3s}). \quad \dots (41)$$

For very large values of R , the only effective contribution to the integral in equation (40) comes from values of p near the point of "stationary phase" at which

$$p = p_0 \text{ and } \tan \chi = \pm \partial k_s / \partial p_0. \quad \dots \dots \dots (42)$$

That is, the only Fourier-Bessel components of the driving function which contribute to the power radiated in directions close to χ are those with radial wave numbers near p_0 defined by equation (42).

Since $0 < \chi < \frac{1}{2}\pi$, equation (42) can be satisfied for only one choice of the ambiguous sign and in the present case where $-k_s(p)$ is the propagation constant of an outward-going wave, we have

$$\begin{aligned} E_{ls}(R \sin \chi, m, R \cos \chi, t) &\sim \frac{\exp[i\theta(p_0)] W_{ls}(p_0)}{\sqrt{(2\pi p_0 R \sin \chi)}} \\ &\times \int_0^\infty dp \exp \left[-i \left\{ \frac{\partial^2 k_s}{\partial p_0^2} \cdot \frac{R \cos \chi}{2} (p - p_0)^2 + 0(p - p_0)^3 \right\} \right], \quad \dots (43) \end{aligned}$$

where

$$\theta(p_0) = -p_0 R \sin \chi - k_s(p_0) R \cos \chi + (m + \frac{1}{2})\pi.$$

The absolute value of the integral in equation (43) does not depend upon the sign of $\partial^2 k_s / \partial p_0^2$, and since

$$\int_{-\infty}^\infty \frac{\exp(-iu)}{2u^{\frac{1}{2}}} du = (\sqrt{\pi}) \exp(-\frac{1}{4}i\pi) = -i \int_{-\infty}^\infty \frac{\exp(iu)}{2u^{\frac{1}{2}}} du, \quad \dots (44)$$

we have from equations (12), (39), and (43)

$$\left. \begin{aligned} \mathbf{E}_s(R, \chi, m, t) &\sim \frac{\exp [i\{\theta(p_0) + \frac{1}{2}\pi + \frac{1}{4}\pi(\partial^2 k_s / \partial p_0^2) / (\partial^2 k_s / \partial p_0^2)\}]}{R \sqrt{\{p_0 \sin \chi \cos \chi \mid \partial^2 k_s / \partial p_0^2\}}} \\ &\quad \times \{-ip_0 E_{1s}, p_0 E_{2s}, E_{3s}\}, \\ \mathbf{H}_s(R, \chi, m, t) &\sim \frac{\exp [i\{\theta(p_0) + \frac{1}{2}\pi + \frac{1}{4}\pi(\partial^2 k_s / \partial p_0^2) / (\partial^2 k_s / \partial p_0^2)\}]}{R \sqrt{\{p_0 \sin \chi \cos \chi \mid \partial^2 k_s / \partial p_0^2\}}} \\ &\quad \times \{-ip_0 H_{1s}, ip_0 H_{2s}, H_{3s}\}, \end{aligned} \right\} \dots (45)$$

where

$$\partial k_s / \partial p_0 = -\tan \chi, \dots (46)$$

provided that*

$$\partial^2 k_s / \partial p_0^2 \neq 0.$$

VI. THE RADIATION FLUX IN A GIVEN DIRECTION

From equation (45), it follows that the mean energy flux associated with the s th mode at a great distance from the source is given by

$$\begin{aligned} \text{Re } \frac{1}{2}(\mathbf{E} \wedge \mathbf{H}^*)_s &= \frac{1}{p_0 R^2 \sin 2\chi \mid \partial^2 k_s / \partial p_0^2 \mid} \\ &\quad \times \text{Re} \{ip_0(E_{2s} H_{3s}^* + E_{3s} H_{2s}^*), ip_0(E_{1s} H_{3s}^* + E_{3s} H_{1s}^*), -p_0^2(E_{1s} H_{2s}^* - E_{2s} H_{1s}^*)\}. \end{aligned} \dots (47)$$

From the vector Maxwell equation

$$-\mu_0 \partial \mathbf{H} / \partial t = \nabla \wedge \mathbf{E},$$

and from equations (36), (37), (38), and (47) we get

$$\begin{aligned} P_s = \text{Re } \frac{1}{2}(\mathbf{E} \wedge \mathbf{H}^*)_s &= \frac{1}{\omega_m \mu_0 p_0 R^2 \sin 2\chi \mid \partial^2 k_s / \partial p_0^2 \mid} \\ &\quad \times \left\{ p_0 \left(p_0^2 E_{2s} E_{2s}^* + \frac{\omega_m^2 - \omega_0^2}{c^2 p_0^2} E_{3s} E_{3s}^* \right), 0, k_s p_0^2 E_{2s} E_{2s}^* + \frac{\omega_m^2 - \omega_0^2}{c^2 p_0^2} \left(\frac{\omega_m^2 - \omega_0^2}{c^2} - p_0^2 \right) \frac{E_{3s} E_{3s}^*}{k_s} \right\} \end{aligned} \dots (48)$$

on using the relation

$$\{(\omega_m^2 - \omega_0^2)/c^2 - p_0^2\} E_{3s}(p, m, z, t) - ik_s p_0^2 E_{1s}(p, m, z, t) = 0, \dots (49)$$

which follows from equation (36).

The power flux $P_{\chi_s} d\chi$ out of an infinite sphere in angle $d\chi$ is given by

$$P_{\chi_s} d\chi = 2\pi R^2 \sin \chi d\chi (P_{rs} \sin \chi + P_{zs} \cos \chi), \dots (50)$$

* Equation (45) is not valid when $\partial^2 k_s / \partial p_0^2 = 0$ but it can be shown that the zeros of $\partial^2 k_s / \partial p_0^2$ form a set of measure zero on the χ -axis, so that they do not contribute to the total radiated power. It remains true, however, that the amplitude of the electromagnetic field and therefore of the radiated power is likely to have a maximum near the zeros of $\partial^2 k_s / \partial p_0^2$.

so that

$$P_{\chi s} d\chi = \frac{\pi d\chi}{\omega_m \mu_0 p_0 \left| \partial^2 k_s / \partial p_0^2 \right|} \times \left\{ p_0^2 E_{2s} E_{2s}^* (p_0 \tan \chi + k_s) + \frac{\omega_m^2 - \omega_0^2}{c^2 p_0^2} E_{3s} E_{3s}^* \left(p_0 \tan \chi + \frac{\omega_m^2 - \omega_0^2 - c^2 p_0^2}{c^2 k_s} \right) \right\}. \quad (51)$$

The total power radiated in the forward direction in both modes is given by

$$P = \sum_{s=1}^2 \int_0^{\frac{1}{2}\pi} P_{\chi s} d\chi. \quad (52)$$

The integral involved in equation (52) is extremely complicated in the general case and one must rely on numerical integration to find a quantitative expression for the radiated power. However, it is possible to draw some useful conclusions about the relative power radiated in different harmonics and different modes from the general nature of the solution. As a preliminary to this discussion, it will be necessary to review briefly some of the properties of the propagation constants $-k_1$, $-k_2$ of the ordinary and extraordinary modes respectively with the aid of the familiar magneto-ionic theory.*

(a) The Propagation Constants

It follows from relation (32) that the propagation constants for a given value of p_0^2 are given by

$$k^2 = \{b \pm \sqrt{(b^2 - 4ad)}\} / 2a, \quad (53)$$

where

$$\left. \begin{aligned} a &= (\omega_m^2 - \omega_0^2) / c^2, \\ b &= \frac{\omega_m^2 - \omega_0^2}{c^2} \left[\frac{2\omega_m^2}{c^2} \left(1 - \frac{\omega_0^2}{\omega_m^2 - \omega_H^2} \right) - p_0^2 \right] - p_0^2 \frac{\omega_m^2}{c^2} \left(1 - \frac{\omega_0^2}{\omega_m^2 - \omega_H^2} \right), \\ d &= \frac{\omega_m^2}{c^2} \left(\frac{\omega_m^2 - \omega_0^2}{c^2} - p_0^2 \right) \left[\frac{\omega_m^2}{c^2} \left(1 - \frac{\omega_0^2}{\omega_m^2 - \omega_H^2} \right)^2 - \frac{\omega_H^2}{c^2} \cdot \frac{\omega_0^4}{\omega_m^2 - \omega_H^2} \right. \\ &\quad \left. - p_0^2 \left(1 - \frac{\omega_0^2}{\omega_m^2 - \omega_H^2} \right) \right]. \end{aligned} \right\} \dots (54)$$

One cannot give a general rule for deciding which choice of the ambiguous sign in equation (53) corresponds to the ordinary and which to the extraordinary wave since this depends upon the relative magnitudes of ω_m^2 , ω_0^2 , and ω_H^2 and each case must be considered separately. In what follows we shall assume that

$$\omega_0^2 < \omega_1^2 < \omega_H^2 < \omega_2^2 - \omega_0^2. \quad (55)$$

* The point of view developed in the present paper is rather different from that normally adopted in magneto-ionic theory but the general survey given by Mitra (1947) contains all the results we need below.

This case is one of the more interesting from the point of view of solar radiophysics since, when it is valid, the ordinary mode can escape from the Sun at the gyro frequency of the rotating fast electron as well as at all the harmonics, while the extraordinary mode can escape at all the harmonics though not, of course, at the fundamental.* It may be noted that the inequality $\omega_1^2 < \omega_H^2$ is a direct consequence of the assumption that the radiating charged particle is an electron rotating under the action of the axial magnetic field in a frame of reference at rest with respect to the background plasma since, in this case,

$$\omega_1^2 = (eB_0/m)^2 = (eB_0/m_0)^2(1 - v_0^2/c^2) = \omega_H^2(1 - v_0^2/c^2) < \omega_H^2,$$

where v_0 is the velocity of the fast particle.

We know from magneto ionic theory that the infinity of the propagation constant at $\omega = \omega_H$ is associated with the extraordinary mode. Now, as ω_m tends to ω_H from below it may be seen from equations (53), (54) that k will tend to infinity when the ambiguous sign is chosen positive. The reverse is true when ω_m tends to ω_H from above, the change of sign arising because the parameter b , defined in equation (54), changes from $+\infty$ to $-\infty$ as ω_m changes from $\omega_H - \varepsilon$ to $\omega_H + \varepsilon$. Accordingly, for radiation at the fundamental of the gyro frequency of the fast electron, for which $\omega_m = \omega_1 < \omega_H$, the positive choice of sign in equation (53) corresponds to the extraordinary mode and vice versa for the higher harmonics of this gyro frequency for which, from the inequality (55), $\omega_m > \omega_H$.

It can be shown that k_1^2 , k_2^2 both decrease monotonically as p increases from zero to the values at which k_1^2 or k_2^2 become zero, which, it may further be shown, coincide with the two zeros of the parameter d defined in equation (54). At these zeros of k_s , $|\partial k_s / \partial p| = \infty$ so that, as one would expect, they correspond to propagation in a direction perpendicular to the axial magnetic field. As p is increased beyond these critical values, k_s^2 becomes negative, so that the associated wave becomes evanescent and a radiation field is produced for only a finite range of values of p which, it can be shown, is defined for the ordinary mode by

$$0 \leq p^2 \leq (\omega_m^2 - \omega_0^2)/c^2 = p_1^2, \quad \dots \dots (56)$$

and for the extraordinary mode by

$$0 \leq p^2 \leq \frac{\omega_m^2}{c^2} \left[\frac{(\omega_m^2 - \omega_H^2 - 2\omega_0^2)\omega_m^2 + \omega_0^4}{\omega_m^2(\omega_m^2 - \omega_H^2 - \omega_0^2)} \right] = p_2^2. \quad \dots \dots (57)$$

(b) Comparison between Radiation in the Ordinary and in the Extraordinary Mode

One important difference between the radiation emitted in the two possible modes is that the ordinary mode is relatively weakly excited for small values of p , that is, in directions making a small angle with the axial magnetic field. To see this we note that

$$pr_0 J'_m(pr_0) = m J_m(pr_0) - pr_0 J_{m+1}(pr_0),$$

* The theory in this paper has been developed assuming a uniform medium but will be valid for the solar corona as long as the macroscopic properties of the medium vary sufficiently slowly with distance.

hence, from equation (36), the amplitudes of the components of the electric field are proportional to

$$\left[k^2 - \frac{\omega_m^2}{c^2} + \frac{\omega_m \omega_0^2}{c^2(\omega_m + \omega_H)} \right] J_m(pr_0) + \text{terms of order } p^2 J_m(pr_0).$$

Neglecting terms of the order of p^2 , it is known that for the ordinary wave

$$k_1^2 - \frac{\omega_m^2}{c^2} + \frac{\omega_m \omega_0^2}{c^2(\omega_m + \omega_H)} = 0,$$

so that the energy carried by this mode is proportional to $p^4 J_m^2(pr_0)$ for sufficiently small values of p . For the extraordinary mode on the other hand

$$k_2^2 - \frac{\omega_m^2}{c^2} + \frac{\omega_m \omega_0^2}{c^2(\omega_m - \omega_H)} = 0,$$

so that the energy carried by this mode is proportional to $J_m^2(pr_0)$ for small values of p .

It can also be shown that the ordinary mode is not excited in a direction perpendicular to the magnetic field. This arises from the fact that the electric vector of the ordinary mode lies parallel to the axial magnetic field in this case and cannot, therefore, be excited by a charged particle of which the velocity and acceleration vectors are both perpendicular to the axial magnetic field.

The fact that the ordinary mode can be excited appreciably over only a limited range of values of the direction angle is one of the reasons for the associated radiation being much less than that in the extraordinary mode; a second and more obvious reason is that the polarization of the extraordinary mode rotates in the same sense as does the radiating electron. This latter fact probably explains why the possibility of the ordinary mode also being generated has been ignored for so long. Furthermore, at the fundamental gyro frequency of the fast electron the relative contribution of the ordinary mode is reduced by an additional effect which arises because the propagation constant of the extraordinary wave near the gyro frequency is very much greater than the propagation constant for the ordinary mode. Thus for the former we have

$$k_2^2 \approx \frac{c^2}{v_0^2} \cdot \frac{\omega_0^2}{\omega_H^2 - \omega_0^2} \left(2 \frac{\omega_H^2 - \omega_0^2}{c^2} - p_0^2 \right), \quad \dots \dots \dots (58)$$

for $\omega_m^2 = \omega_H^2(1 - v_0^2/c^2)$ and for values of p_0^2 which obey the inequality (56), at least as long as

$$v_0^2/c^2 \ll \omega_0^2/(\omega_H^2 - \omega_0^2).$$

The propagation constant for the ordinary mode on the other hand is nearer to the free space value, being given by

$$k_1^2 \sim \frac{\{(2\omega_H^2 - \omega_0^2)/c^2 - p_0^2\} \{(\omega_H^2 - \omega_0^2)/(c^2 - p_0^2)\}}{2(\omega_H^2 - \omega_0^2)/c^2 - p_0^2}. \quad \dots \dots (59)$$

The characteristic admittance of the extraordinary wave is therefore much greater than that of the ordinary wave with which it is effectively in parallel

from the point of view of the source current generator, so that almost all the available power will be generated in the extraordinary mode.

To make this argument more concrete, we note from the relations (58), (59) that the average value of p_0^2 is of the order of magnitude of ω_H^2/c^2 , for radiation at the fundamental gyro frequency of the rotating electron, so that the parameter pr_0 which occurs as the argument of the Bessel function in equation (36) is of the order of magnitude of v_0/c . As long as $v_0/c \ll 1$, we have

$$J_1(pr_0) \simeq pr_0 J_1'(pr_0) \simeq \frac{1}{2} v_0/c, \quad \dots \quad (60)$$

while

$$J_1(pr_0) - pr_0 J_1'(pr_0) = pr_0 J_2(pr_0) \simeq \frac{1}{8} (v_0/c)^3. \quad \dots \quad (61)$$

Now for the large values of k_2 we see from equation (51) that the power radiated in the χ -direction in the extraordinary mode is of the order of magnitude of

$$(\omega_H^2/c^2) E_{22} E_{22}^* k_2 + E_{32} E_{32}^* \omega_H/c.$$

From equation (36), (58), (60), and (61), we see that for the extraordinary mode the term $E_{22} E_{22}^* k_2 \omega_H^2/c^2$ is of the order of magnitude of v_0^3/c^3 , while the term $E_{32} E_{32}^* \omega_H/c$ is of the order of magnitude of v_0^2/c^2 . Hence, just as in the free space case, the total power radiated in the extraordinary mode at the first harmonic of the gyro frequency of the fast electron is proportional to v_0^2/c^2 . For the ordinary mode, on the other hand, in which k_1 is also of the order of magnitude of ω_H/c , we see from the same equations that E_{21} , E_{31} are both of the order of magnitude of v_0^3/c^3 . The total power radiated in the ordinary mode at the first harmonic of the gyro frequency of the fast electron is therefore of the order of magnitude of v_0^6/c^6 .

If we now consider the m th harmonic of the gyro frequency, however, at which p and also k_1 , k_2 are all of the order of magnitude of $m\omega_H/c$, it follows, from equation (36) and the fact that

$$J_m(pr_0) \simeq (m^m/m! 2^m) (v_0/c)^m \quad \text{for } mv_0/c \ll 1,$$

that the power radiated in either mode is proportional, for small enough values of v_0/c , to $(v_0/c)^{2m}$. In particular, it follows that the power radiated in the ordinary mode at the second harmonic of the gyro frequency of the fast electron is larger by a factor $(c/v_0)^2$ than that radiated at the fundamental frequency in the same mode. Finally, it may be noted that the power radiated in the ordinary mode at the fundamental varies in the same way with v_0/c as the power radiated in this extraordinary mode at the third harmonic, but because the latter mode should be generated more efficiently one would expect that it would carry appreciably more power. This is confirmed by the numerical results given below.

(c) Numerical Results

The calculations shown in Table 1 below were carried out for the special case

$$\beta^2 = v_0^2/c^2 = 0.1, \quad \dots \quad (62)$$

corresponding to a kinetic energy of the rotating fast electron of approximately 27.7 keV. In order to bring out the effect of the background plasma electrons

we have compared the powers radiated in different modes and at different harmonics in the two special cases, (a) $\omega_0^2/\omega_H^2=0.8$, (b) $\omega_0^2/\omega_H^2=0$. To show how the radiated power varies with the velocity of the fast electron we have expressed the total power radiated at the m th harmonic as a numerical quantity times β^{2m} . The results are exact only for the special case of equation (62), but they will be accurate to better than 10 per cent. for values of β^2 in the range

$$0 < \beta^2 < 0.15.$$

As might be expected, the effect of the plasma electrons is most serious at the fundamental frequency. In every case, however, it may be seen that the power radiated is lower than in the free space case.

TABLE 1

COMPARISON BETWEEN POWERS RADIATED IN DIFFERENT MODES AND AT DIFFERENT HARMONICS FOR TWO VALUES OF THE DENSITY OF THE BACKGROUND PLASMA ELECTRONS

Harmonic	Total Power* Radiated in Forward Direction	Percentage of Power Radiated in Forward Direction in Ordinary Mode	Total Power* Radiated in Forward Direction	Percentage of Power Radiated in Forward Direction in Ordinary Mode
1	$0.0755\beta^2(1-\beta^2)C$	$\beta^4 \times 1.53\%$	$0.322\beta^2(1-\beta^2)C$	$\beta^4 \times 1.63\%$
2	$0.59\beta^4(1-\beta^2)C$	2.1%	$0.729\beta^4(1-\beta^2)C$	2.5%
3	$1.18\beta^6(1-\beta^2)C$	6.5%	$1.68\beta^6(1-\beta^2)C$	4.4%

* The expressions for the power are in watts when $C=2.48 \times 10^{-22}H^2$ and H is in gauss.

Besides changing the total radiated power, the plasma electrons also modify appreciably the polar diagram, especially at the fundamental frequency. This effect is illustrated in Figures 1 (a) and 1 (b), where we have plotted the polar diagrams for the radiation at the first three harmonics in the ordinary mode for the case (a) $\omega_0^2/\omega_H^2=0.8$, (b) $\omega_0^2/\omega_H^2=0$. In Figures 2 (a) and 2 (b), we have plotted polar diagrams for the radiation, in the extraordinary mode, into the second and third harmonics: the polar diagram for the first harmonic in the ionized plasma has not been plotted since at this frequency the radiation cannot escape from the Sun.

VII. DISCUSSION AND CONCLUSIONS

The main result of this paper is to confirm that a fast electron rotating with velocity $v_0 \ll c$ in an ionized plasma permeated by a uniform magnetic field will radiate energy both in the ordinary and in the extraordinary modes, but predominantly in the latter. Thus the power radiated at the fundamental gyro frequency in the ordinary mode is proportional to $(v_0/c)^6$ in contrast with the extraordinary mode in which the power radiated is proportional to $(v_0/c)^2$; indeed the first harmonic power in the ordinary mode is appreciably less than

that radiated in the third harmonic—at least in the numerical cases considered in this paper. At the m th harmonic, $m > 1$, the power is still radiated predominantly in the extraordinary mode, though for both modes it is proportional to $(v_0/c)^{2m}$ as long as $(v_0/c)^2$ is sufficiently small. These conclusions are valid

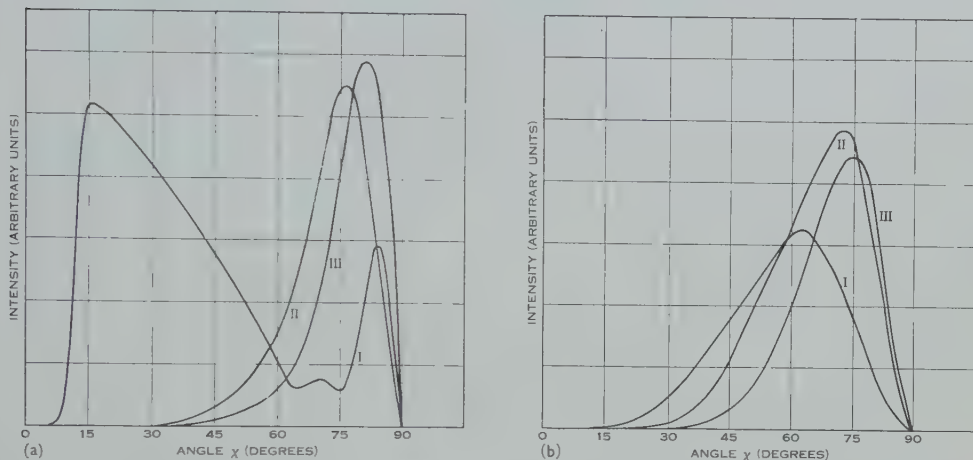


Fig. 1.—Emission polar diagrams for the ordinary mode. (a) With $\omega_0^2/\omega_H^2 = 0.8$; (b) for free space, $\omega_0^2/\omega_H^2 = 0$. The emission per unit solid angle is shown as a function of χ , the inclination to the magnetic field. The intensity scale is arbitrary and is different for each curve. Curve I fundamental; curve II, second harmonic; curve III, third harmonic.

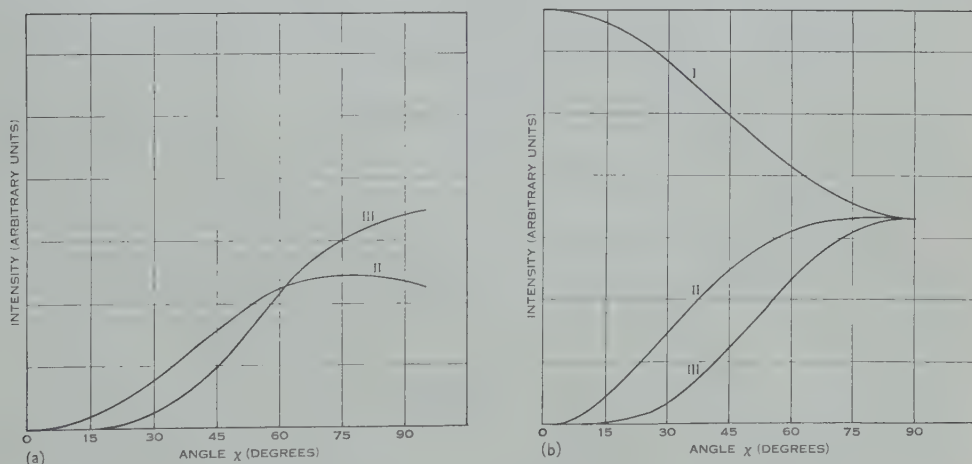


Fig. 2.—Corresponding emission polar diagrams for the extraordinary mode.

not only in media so dense that the plasma frequency is comparable with the gyro frequency but also in the limiting case of free space propagation in which the plasma frequency tends to zero, as shown in Appendix I.

To see the relevance of these results to gyro-type theories for the non-thermal solar radiation, let us first consider the bursts of spectral types II and

III. In these phenomena, the radiation at a given instant is concentrated into one or two frequency bands, the centre frequency of which decreases with time; very rapidly for the type III bursts, less rapidly for the type II bursts (Wild, Murray, and Rowe 1954). Recent observations have shown that in the case of type III bursts this decrease of frequency is associated with a rapid movement of the radiating disturbance (Wild, personal communication). In those cases in which two frequency bands are observed the data are consistent with the belief that the higher frequency band is simply the second harmonic of the lower. There is virtually no evidence that the third or higher harmonics are radiated; if these are present then their intensity, in the great majority of cases, must be at least one order of magnitude lower than that in the first two harmonics. The apparent dimensions of the disturbed regions in a type III burst are large, being up to 10 per cent. or even more of the area of the visible solar disk (Wild and Sheridan 1958). In general, these bursts are unpolarized; however, in quite a number of cases there is an appreciable component of circular polarization, and when this is so the sense of polarization is the same in both the high and the low frequency bands (Komesaroff 1958).

If these features are to be explained on the gyro theory, one has to imagine either that a burst of fast electrons is fired out from the Sun or that an electromagnetic disturbance travels outward giving rotational energy to a fraction of the local plasma electrons as it goes. On either picture the decrease in frequency with time would be attributed to a decrease in gyro frequency with increasing height above the solar corona. The energy in the fundamental band would have to be carried entirely by the ordinary mode, since the extraordinary mode cannot escape at the fundamental gyro frequency (Ryle 1950). On the other hand, one would expect the energy at the higher harmonics to be carried predominantly by the extraordinary mode for the reasons discussed in this paper. One might hope to explain the fact that the majority of bursts are not strongly polarized by the large size of the radiating regions together with the assumption that the sense of the magnetic field is different in different parts of the radiation front. However, when the bursts were polarized, one would expect the sense of polarization to be different in the higher and lower frequency bands and this is in disagreement with the few observations so far available (Komesaroff 1958). It is also very hard to see why the third harmonic should not be present and stronger than the first since, as we have shown, the power radiated at the fundamental frequency in the ordinary mode is less than that radiated at the third harmonic in the ordinary mode,* and very much less than that radiated at the third harmonic in the extraordinary mode. A similar objection is that the power radiated at the fundamental gyro frequency in the ordinary mode is very much smaller than the power radiated either in the extraordinary or in the ordinary mode at the second harmonic, and this is in disagreement with the observed ratio of the power radiated in the upper and lower frequency bands; at least

* The possibility that the absence of the third harmonic could be due to differential absorption can be eliminated, since the ordinary mode at the fundamental gyro frequency will be more heavily absorbed than either mode at the third harmonic of the gyro frequency, especially in the case when the plasma and gyro frequencies are of comparable magnitude.

in some cases. Once again the discrepancy is too great to be explained away by differential absorption.

Admittedly, we assumed in our quantitative analysis that the velocity v_0 of the fast radiating electrons was significantly less than c , but, if this were not the case, appreciable power would be radiated at quite a few harmonics of the gyro frequency, which is contrary to observation.

A final objection to the gyro theory of the rapid drift bursts is that the first harmonic can escape only when the plasma frequency is less than the gyro frequency. Since the type III bursts at least are often associated with quite small sunspots this requirement is not very likely to be met.

It is possible that some of these objections would be removed or at least mitigated by a more exact analysis in which one allows for the finite temperature of the background plasma and for Doppler effects in the radiation from the gyrating electrons. However, at the present time the weight of evidence is still strongly against the gyro theory, at least as far as the types II and III bursts are concerned.

The position is perhaps more hopeful in the case of the type I disturbances, which are strongly associated with large sunspots and which consist of short-duration bursts, with a narrow frequency spread, superimposed upon a continuous disturbance of much longer duration and with a wide frequency spread. Both bursts and continuum are strongly circularly polarized and it is conceivable that the continuum disturbance may actually consist of a very large number of small bursts occurring over a wide range of heights in the solar corona. There is some rather weak evidence of correlation between bursts at different frequencies, but there is definitely no sign of a 2 to 1 frequency correlation as in types II and III.

Much of the data is compatible with the suggestion that the type I disturbances are due to second harmonic gyro radiation. However, one objection to this is that such an explanation would lead one to expect that the radiation would be in the extraordinary mode: the evidence on this point is too scanty to be conclusive, but it suggests that the radiation is in the ordinary mode, at least in the majority of cases (Payne-Scott and Little 1952). If this is confirmed, it would seem that the gyro hypothesis could be saved only if (a) the absorption in the extraordinary mode were greater by about three orders of magnitude than that in the ordinary mode at the same frequency* or if (b) the plasma frequency at the point of generation were so high, at least in the majority of cases, that the second harmonic could not escape in the extraordinary mode. It may be noted that, in a zero temperature plasma, the conditions that the ordinary mode at the m th harmonic should escape and that the extraordinary mode should be trapped is

$$\omega_0 < m\omega_H < \sqrt{\{m/(m-1)\}}\omega_0.$$

* This would be quite impossible, in any conceivable model of the solar corona, if absorption were due simply to free-free transitions in the background plasma. However, as has been pointed out by Gross (1951), a finite temperature plasma permeated with a magnetic field may possess stop bands at the harmonics of the gyro frequency and it is not inconceivable that this phenomenon could discriminate against the extraordinary mode.

Here ω_0 , ω_H are the plasma and gyro angular frequencies of the plasma, and we have assumed $v_0^2/c^2 \ll 1$. It is not implausible that this condition might be met, when $m=2$, above a large sunspot.

Of course, it is not sufficient merely to prove that waves with the right sense of circular polarization could both be generated and escape; one also has to explain quantitatively the observed energy flux. Now in the larger type I disturbances the effective black-body temperature of the source may be as high as 3×10^{10} °K or even higher (Wild and Sheridan 1958), and the electrons of a gas at this temperature would have an average energy of $\sim 3 \times 10^6$ eV. However, the kinetic energy of the gyrating electrons must be very much less than this or the bursts would not consist of isolated narrow bands, nor would the emergent radiation be so strongly circularly polarized. To obtain a radiation flux with a black-body temperature much greater than that appropriate to the average kinetic energy of a non-equilibrium electron gas implies a process of radiation transfer very different from that usually encountered in stellar atmospheres, but as pointed out elsewhere (Twiss, in preparation) it is conceivable that conditions could arise in the disturbed solar corona where the absorption coefficient was negative so that the medium behaved like an amplifier. If this latter effect were really present one could certainly explain the observed radiation flux. However, a detailed discussion of this possibility involves a theory for the case where a large number of gyrating electrons is present, and this is beyond the scope of the present paper.

VIII. ACKNOWLEDGMENTS

The authors wish to express their thanks to Mrs. Marjorie Gorzkos, who carried out most of the numerical computation, and to Mr. J. P. Wild for helpful discussions during the preparation of the paper.

IX. REFERENCES

- GROSS, E. P. (1951).—*Phys. Rev.* **82**: 232–42.
 KIEPENHEUER, K. O. (1946).—*Nature* **158**: 340.
 KOMESAROFF, M. (1958).—*Aust. J. Phys.* **11**: 201–14.
 KRUSE, U. E., MARSHALL, L., and PLATT, J. R. (1956).—*Astrophys. J.* **124**: 601–4.
 MITRA, S. K. (1947).—“The Upper Atmosphere.” (Roy. Asiatic Soc. of Bengal: Calcutta.)
 PAYNE SCOTT, RUBY, and LITTLE, A. G. (1952).—*Aust. J. Sci. Res. A* **5**: 32–49.
 RYLE, M. (1950).—*Rep. Progr. Phys.* **13**: 184–246.
 SCHOTT, G. A. (1912).—“Electromagnetic Radiation.” (Cambridge Univ. Press.)
 TWISS, R. Q. (1952).—*Phys. Rev.* **88**: 1392–407.
 WILD, J. P., MURRAY, J. D., and ROWE, W. C. (1954).—*Aust. J. Phys.* **7**: 439–59.
 WILD, J. P., and SHERIDAN, K. V. (1958).—*Proc. Inst. Radio Engrs., N.Y.* **46**: 160–71.

APPENDIX I

Resolution of the Radiated Field into Ordinary and Extraordinary Modes in the Limit of Vanishing Electron Density

Even when the density of the background plasma electrons is vanishingly small, the magneto-ionic modes still have defined polarization so that the radiation emitted by an electron gyrating "*in vacuo*" under the action of an external magnetic field may still be resolved into ordinary and extraordinary modes.

In spherical coordinates (r, χ, φ) the magneto-ionic modes in the limiting case of zero plasma frequency obey the constraint (Mittra 1947)

$$E_\varphi/E_\chi = -H_\chi/H_\varphi = (-i/2 \cos \chi)[(\omega_H/\omega_m) \sin^2 \chi \pm \{(\omega_H/\omega_m)^2 \sin^4 \chi + 4 \cos^2 \chi\}^{\frac{1}{2}}], \quad \dots\dots\dots (\text{A1})$$

where ω_H is the angular gyro frequency of a background plasma electron,
 ω_m is the m th harmonic of the angular gyro frequency of the fast gyrating electron,

and where the upper sign refers to the extraordinary, the lower sign to the ordinary mode.

For the radiation at the fundamental gyro frequency of the fast electron

$$\omega_H/\omega_m = \omega_H/\omega_1 = (1 - \beta^2)^{-\frac{1}{2}} \simeq 1 + \frac{1}{2}\beta^2$$

to the first order in $\beta^2 = v_0^2/c^2$, where v_0 is the velocity of the fast electron.

Accordingly, at the first harmonic, we have

$$(E_\varphi/E_\chi)_e = [(E_\varphi/E_\chi)_o]^{-1} \simeq -i \cos \chi \{1 + \frac{1}{2}\beta^2 \sin^2 \chi / (1 + \cos^2 \chi)\}, \quad \dots (\text{A2})$$

where the subscripts e and o refer to the extraordinary and to the ordinary mode respectively.

The polarization ellipses for the two modes are similar, but their major axes are mutually perpendicular.

Now for the radiated field (Schott 1912) we have

$$\left(\frac{E_\varphi}{E_\chi}\right)_{\text{rad}} = \frac{-i}{\cos \chi} \cdot \frac{J_{m-1}(m\beta \sin \chi) - J_{m+1}(m\beta \sin \chi)}{J_{m-1}(m\beta \sin \chi) + J_{m+1}(m\beta \sin \chi)} \quad \dots (\text{A3})$$

Hence for the fundamental mode, $m=1$, we have, to the first order in β^2 ,

$$(E_\varphi/E_\chi)_{\text{rad}} = (-i/\cos \chi)(1 - \frac{1}{4}\beta^2 \sin^2 \chi). \quad \dots\dots\dots (\text{A4})$$

Comparison between equations (A2) and (A4) shows that, although the radiated power is carried predominantly by the extraordinary mode, the ordinary mode is also excited. To determine the proportion of ordinary mode excited at a given angle χ to the axial magnetic field, let $i\rho^{\frac{1}{2}}$ be the ratio of the major axis of the polarization ellipse of the ordinary mode to that of the extraordinary mode. Then we have

$$\begin{aligned} E_{\chi e} &= E_0 \exp(i\omega t), & E_{\varphi e} &= (-i/\cos \chi)\{1 + \frac{1}{2}\beta^2 \sin^2 \chi / (1 + \cos^2 \chi)\}E_0 \exp(i\omega t), \\ E_{\chi o} &= (-i/\cos \chi)\{1 + \frac{1}{2}\beta^2 \sin^2 \chi / (1 + \cos^2 \chi)\}i\rho^{\frac{1}{2}}E_0 \exp(i\omega t), & E_{\varphi o} &= i\rho^{\frac{1}{2}}E_0 \exp(i\omega t), \end{aligned} \quad \dots\dots\dots (\text{A5})$$

where E_0 is a constant of proportionality.

Hence

$$\frac{\rho_1^{\frac{1}{2}} \sec \chi \{1 + \frac{1}{2} \beta^2 \sin^2 \chi / (1 + \cos^2 \chi)\}}{1 - \rho_1^{\frac{1}{2}} \sec \chi \{1 + \frac{1}{2} \beta^2 \sin^2 \chi / (1 + \cos^2 \chi)\}} = \frac{-i}{\cos \chi} (1 - \frac{1}{4} \beta^2 \sin^2 \chi),$$

so that

$$\rho_1 = \frac{1}{16} \beta^4 [\sin^4 \chi \cos^2 \chi (3 + \cos^2 \chi)^2 / (1 + \cos^2 \chi)^4], \quad \dots \quad (\text{A6})$$

and the ratio of the *power* radiated in the ordinary mode to that in the extraordinary mode at the fundamental gyro frequency of the fast electron is therefore proportional to β^4 as in the case, considered in the text, in which the plasma frequency associated with the background electrons is comparable with the gyro frequency.

For harmonics higher than the first, it is sufficient to work to zero order in β^2 since, for the magneto-ionic modes at the m th harmonic, $m > 1$,

$$E_\phi / E_\chi \simeq -i \{ (\frac{1}{2} m) \sin \chi \tan \chi \pm [1 + \{ (\frac{1}{2} m) \sin \chi \tan \chi \}^2]^{\frac{1}{2}} \}, \quad \dots \quad (\text{A7})$$

where, as before, the positive and negative choices of sign in equation (A7), refer to the extraordinary and to the ordinary mode respectively.

For the radiated field, on the other hand, we have from equation (A3)

$$(E_\phi / E_\chi)_{\text{rad}} = -i \sec \chi,$$

if we neglect terms $O(\beta^2)$.

Hence

$$\frac{\rho_m^{\frac{1}{2}} - \{ (\sin \chi \tan \chi) / 2m + [1 + \{ (\sin \chi \tan \chi) / 2m \}^2]^{\frac{1}{2}} \}}{1 + \rho_m^{\frac{1}{2}} \{ (\sin \chi \tan \chi) / 2m + [1 + \{ (\sin \chi \tan \chi) / 2m \}^2]^{\frac{1}{2}} \}} = -\sec \chi,$$

and

$$\rho_m = \frac{1 - \cos \chi \{ (\sin \chi \tan \chi) / 2m + [1 + \{ (\sin \chi \tan \chi) / 2m \}^2]^{\frac{1}{2}} \}}{\cos \chi + \{ (\sin \chi \tan \chi) / 2m + [1 + \{ (\sin \chi \tan \chi) / 2m \}^2]^{\frac{1}{2}} \}}, \quad \dots \quad (\text{A8})$$

which is independent of β^2 .

SHORT COMMUNICATIONS

FERROMAGNETIC EXCHANGE BETWEEN COUPLED PAIRS OF ELECTRONS*

By F. D. STACEY†

Incomplete shells of $3d$ electrons in the first transition series of elements are responsible for the ferromagnetism of these elements, but the parallel coupling of the spins of these electrons remains an assumption in the theory of ferromagnetism. It is required that exchange interaction between the $3d$ electrons of neighbouring atoms should be positive (the sign causing parallel alignment of spins). Mathematically, treatment of ferromagnetic exchange is similar to that of the homopolar chemical bond, in which it is clear that exchange interaction is negative, giving a lower energy for antiparallel spins than for parallel spins. Mathematical difficulties have prevented any adequate demonstration that the reverse is the case in ferromagnetism.

Zener and Heikes (1953) maintained that exchange interactions are always negative, and developed a theory in which parallel alignment of $3d$ electrons occurs by exchange with the conduction electrons. There are some serious objections to this theory and it is more generally assumed that direct coupling between $3d$ electrons is responsible for their parallel spin alignment (Van Vleck 1953). The object of the present note is to point out that the fundamental objection to positive exchange by direct interaction is removed if the elementary magnets are parallel coupled spins which interact in pairs. This requires that each magnetically active atom has two positive holes in its $3d$ shell, with spins coupled parallel by Hund's rule, a suggestion first made by Mott and Jones (1936) for essentially experimental reasons.

The wave function of two electrons, numbered 1 and 2, associated with nuclei a and b is written :

$$\text{antiparallel spins : } \psi_n = (\psi_a(1)\psi_b(2) + \psi_a(2)\psi_b(1))(\alpha(1)\beta(2) - \alpha(2)\beta(1))$$

$$\text{parallel spins : } \psi_p = (\psi_a(1)\psi_b(2) - \psi_a(2)\psi_b(1)) \begin{cases} \alpha(1)\beta(2) + \alpha(2)\beta(1) \\ \text{or } \alpha(1)\alpha(2) \\ \text{or } \beta(1)\beta(2) \end{cases}$$

In each wave function the first bracket represents the orbital part and the second bracket the spin part, in which α and β define opposite spin directions. The functions are constructed to be antisymmetric with respect to exchange of the two electrons, as required for the treatment of electrons as Fermi particles.

* Manuscript received March 4, 1958.

† Department of Geophysics, Australian National University, Canberra.

In the hydrogen molecule problem, ψ_n^2 and ψ_p^2 are multiplied by the Coulomb interaction potential for two nuclei and two electrons and integrated over the volume occupied by ψ_n and ψ_p . (The detailed explanation of this is given by Heitler (1944)). The resulting energy is lower for ψ_n than ψ_p , indicating the negative exchange which causes homopolar bonding.

If the geometry of the wave functions remains essentially unaltered, but 1 and 2 represent parallel coupled pairs instead of individual electrons, and if the coupling within pairs is much stronger than between pairs, we may consider the exchange of pairs and not individual electrons. In this case the pairs behave as bosons rather than fermions, since the wave function of four electrons is symmetric with respect to exchange of pairs, and we must interchange the spin parts of wave functions ψ_n and ψ_p . ψ_n will still give the lower energy but it now gives parallel alignment of all four spins.

The spin wave functions will necessarily be more complicated than has been indicated because the moment of pair 2 in the direction of 1 may be zero as well as ± 2 Bohr magnetons. However, this modification does not need to be given detailed consideration since Van Vleck (1945) established that exchange energy is proportional to the scalar product of spin vectors.

An obvious implication of the present argument is that the existence of spins coupled in even numbers is a necessary prerequisite for ferromagnetism. Mott and Jones (1936) pointed out that there was no objection to this for materials with much less than 2 Bohr magnetons (μ_B) per atom. Materials with 2.5 – $3.5 \mu_B$ would contradict the coupled-spin rule, but in the manganese and chromium alloys and compounds listed by Bozorth (1951) the only μ_B numbers much greater than two are nearly four. This gives an immediate explanation for the absence of ferromagnetism in pure manganese, in which the uncompensated spins are probably coupled mainly in threes instead of twos and fours, without postulating a special dependence of exchange energy on atomic spacing. It also allows for the depression of the μ_B number in iron to nearly two, on the Heitler-London-Heisenberg model, without appealing to the band theory of metals to explain the reversed spins.

Metallic chromium provides a contradiction since this could be ferromagnetic with approximately four spins per atom, and gadolinium, with almost exactly seven spins per atom, would still require an indirect exchange mechanism.

The author wishes to acknowledge the helpfulness of a discussion with Professor K. J. Le Couteur.

References

- BOZORTH, R. M. (1951).—"Ferromagnetism." p. 334. (D. Van Nostrand: New York.)
 HEITLER, W. (1944).—"Elementary Wave Mechanics." (Oxford Univ. Press.)
 MOTT, N. F., and JONES, H. (1936).—"Theory of the Properties of Metals and Alloys." p. 224. (Oxford Univ. Press.)
 VAN VLECK, J. H. (1945).—*Rev. Mod. Phys.* **17**: 27.
 VAN VLECK, J. H. (1953).—*Rev. Mod. Phys.* **25**: 220.
 ZENER, C., and HEIKES, R. R. (1953).—*Rev. Mod. Phys.* **25**: 191.

THE SOFT X-RAY L_{23} EMISSION SPECTRUM OF MAGNESIUM FROM SOLID AND EVAPORATED TARGETS*

By R. S. CRISP†

Spectroscopic studies in the soft X-ray region of 50–1000 Å have used both photographic and photomultiplier detection of the radiation diffracted in the grating spectrograph (see, for example, Tomboulion 1957). The photomultiplier method offers the advantage that it is possible to record the time variation in the intensity of radiation emitted at a particular wavelength as changes proceed in the condition of the target.

Using a 1 m grazing-incidence spectrograph with photomultiplier detector (Fisher, Crisp, and Williams 1958) a detailed study has been made of the magnesium L_{23} emission band at 250 Å.

Experimental Method and Results

The targets, which were water cooled, consisted either of a solid magnesium specimen scraped clean *in vacuo*, or a layer of magnesium evaporated onto a solid copper backing. The variation with time of the peak intensity of the band at 250 Å is illustrated in Figure 1, in which curve (a) is for solid magnesium and (b) and (c) are for thick and thin films of magnesium evaporated onto copper. The zero of the time scale corresponds to switching the electron beam onto the freshly prepared surface. The initial small drop in (a) is attributed to a rapid oxidation of the freshly cut metal surface. This effect has also been observed for solid targets of the alkali metals where there is a much greater fall to about 30 per cent. of the initial intensity in some 60 sec. The decrease is approximately exponential, in agreement with the findings of Kingston (1951) for potassium and calcium.

The initial fall is not observed in (b) because there is about 15 sec delay between ceasing evaporation and turning on the electron beam and presumably there is a more rapid oxidation of evaporated magnesium. The initial rise in (c) is attributed to homogenization of the thin film under the action of the electron beam.

The linear fall in each case is caused by the deposition of carbon from residual organic vapours when bombarded by the electron beam (Ennos 1953, 1954; Fisher, Crisp, and Williams 1958). After 35 or 40 min bombardment a brownish film is clearly visible on the target surfaces. Contamination curves of the above kind were recorded at the peak, dip, and hump at 250, 253, and 263 Å respectively

* Manuscript received April 8, 1958.

† Department of Physics, University of Western Australia, Nedlands, W.A.

(P , D , H in Fig. 2) and were found to be the same for any one type of target. From this it is concluded that the relative intensities of these features do not alter as contamination effects proceed. This makes possible a correction when comparing intensities observed at different wavelengths at different times.

A large number of spectra were recorded for both solid and evaporated targets, the interval after preparation of the target surface being sufficient to

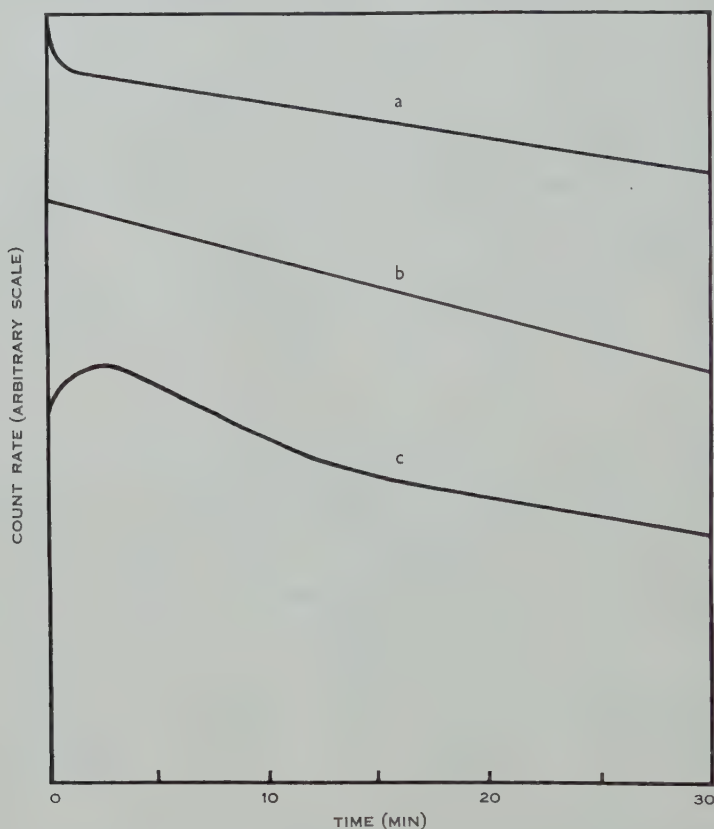


Fig. 1.—Typical curves illustrating the variation of the intensity of radiation emitted at a particular wavelength with time for a freshly prepared target surface at 8×10^{-6} mm Hg. (a) Solid magnesium, scraped surface; (b) thick evaporated magnesium film, evaporating furnace run for about 15 sec; (c) thin evaporated magnesium film, evaporating furnace run for about 5 sec.

ensure that the contamination rate was linear. Scanning speeds were selected with due consideration of the response time of the recording devices (Fisher, Crisp, and Williams 1958). Counting losses were negligible for the counting rates and the resolving time used.

A typical record for a solid target is shown in Figure 2. R indicates the reversal of the direction of traverse. Although the speed of traverse has also

changed at R the abscissae still correspond to a linear time scale and therefore the correction $\pm\delta$ to be applied to the peak intensity P when comparing it with the hump intensity H is valid. A similar correction was applied to the dip intensity D and values were calculated for the ratios $(P-B)/(H-B)$ and $(D-B)/(H-B)$, where B is the background. The mean values of these ratios agreed for spectra from both types of target to within the probable error of 3 per cent. It is therefore concluded that the band shapes are constant, at least in so far as they are represented by these ratios.

Measurements were also made of the edge breadth and wavelength as defined by Skinner (1940). The edge breadths for solid and evaporated targets are constant to within 0.02 eV and the edge wavelengths to within 0.09 Å. Because of the different thermal conductivities of the solid magnesium and copper targets

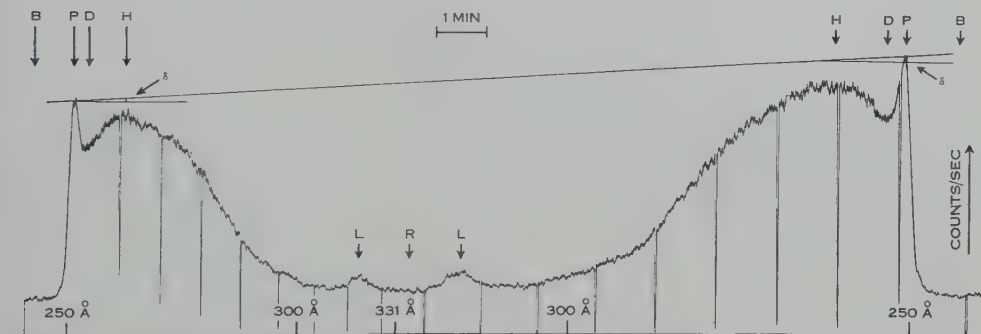


Fig. 2.—Reproduction of chart record for the L_{23} band emitted from a solid magnesium target. Target voltage 4 kV. Target current 4 mA. Maximum counting rate 4000 counts/sec. For explanation of lettering see text. L indicates the (L_3-L_1) line identified by Tombouliau and Cady (1941).

an estimate was made of the difference in surface temperature resulting from the electron beam. This came to about 60 °C and allowed a small correction to be made to the observed edge breadths. No density of states curve has been deduced from the present data, since the optical resolution of 0.9 Å causes considerable loss in the observed height of the sharp peak. It is of interest to notice, however, that in the present instrument, where the plane of the analyser slit lies along a radius of the Rowland circle, the experimental "window" $d\lambda$ Å wide is constant for all wavelengths and the intensities recorded are certainly proportional to the number of photons in the wavelength range $d\lambda$ Å (Fisher 1954). There is thus no doubt what power of the frequency of radiation to use when translating from observed spectra to density of states curves (Tombouliau 1957), in contrast to the photographic case.

The author wishes to acknowledge most gratefully the guidance, help, and suggestions of Professor C. J. Birkett Clews and Dr. S. E. Williams, and to thank the Australian Atomic Energy Commission for a Post-graduate Research Studentship during the tenure of which this work was carried out.

References

- ENNOS, A. E. (1953).—*Brit. J. Appl. Phys.* **4**: 101.
 ENNOS, A. E. (1954).—*Brit. J. Appl. Phys.* **5**: 27.
 FISHER, P. (1954).—*J. Opt. Soc. Amer.* **44**: 665.
 FISHER, P., CRISP, R. S., and WILLIAMS, S. E. (1958).—*Optica Acta*. (In press.)
 KINGSTON, R. H. (1951).—*Phys. Rev.* **84**: 944.
 SKINNER, H. W. B. (1940).—*Phil. Trans. A* **239**: 95.
 TOMBOULIAN, D. H. (1957).—"Handbuch der Physik." Vol. 30, p. 246. (Springer: Berlin.)
 TOMBOULIAN, D. H., and CADY, W. M. (1941).—*Phys. Rev.* **59**: 422.

FREEZING NUCLEUS MEASUREMENTS IN JANUARY 1957*

By E. G. BOWEN†

It is well known that the appearance of the ice phase in supercooled clouds is one of the more important phenomena leading to the formation of rain. The formation of ice crystals is itself dependent on the freezing nucleus content of the atmosphere. Little is known about the nature and origin of natural freezing nuclei, but daily measurements of freezing nucleus concentration made in different parts of the world during the month of January in the three years 1954, 1955, and 1956 have suggested that some fraction of them might be of extraterrestrial origin (Bowen 1956). The present note describes a further series of measurements made during January 1957.

If the nuclei were of extraterrestrial origin, one might predict that:

- (1) there would be three peaks in the freezing nucleus concentration in the period from January 10 to February 1;
- (2) the peaks would occur on approximately the same dates in the northern and southern hemispheres;
- (3) they would occur on approximately the same dates as in previous years, namely, January 13, 22, and 30.

It is natural to expect that the date of these occurrences would show a scatter, and in previous years this has averaged ± 2 days. In any one place and in any given year it could obviously exceed this value. The spatial distribution of nuclei through the atmosphere is unknown but it is almost certain to be patchy. All the peaks would not necessarily appear in all localities, therefore, but in previous years they have appeared in 19 out of 27, that is, approximately 70 per cent. of the predicted occasions.

In this particular series of observations, measurements were made at one station only in 1954, at four stations in widely separated parts of the world in

* Manuscript received March 24, 1958.

† Division of Radiophysics, C.S.I.R.O., University Grounds, Chippendale, N.S.W.

1955, and again at four stations in 1956. The 1957 observations were made at eight different stations, six of which were in or around the continent of Australia and two in the U.S.A. These measurements were made at ground level using an expansion type cold chamber designed by Bigg and Warner (Warner 1957). It is known from previous experience that great care must be exercised in the

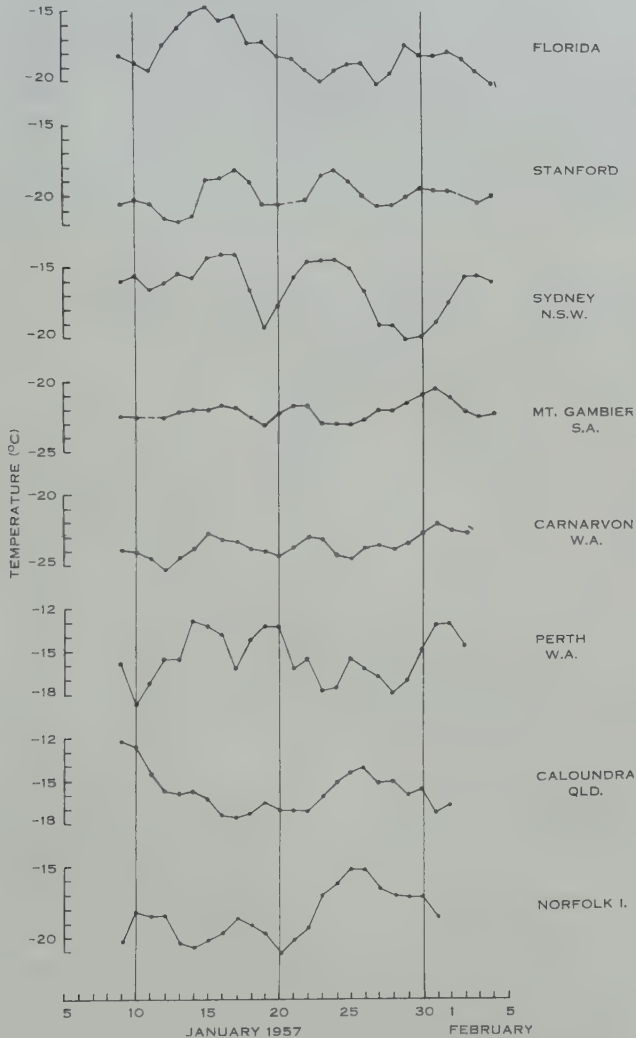


Fig. 1.—Temperatures at which 0.1 freezing nuclei per litre were observed at eight different stations during January 1957.

choice of site so as to be free from the effects of contamination. Broadly speaking, difficulties may arise from: (a) terrestrial dust of local origin; (b) industrial smoke, which in some cases is known to enhance the nucleus count and in others is suspected of depressing it; (c) widespread rain, which might wash the nuclei out of the atmosphere; (d) frost or snow particles from trees and vegetation when the temperature is near freezing.

The sites used in 1957 were on or near the sea coast and satisfied most of the above conditions as far as possible.

Results

Figure 1 gives the temperature at which a concentration of 0.1 nuclei per litre was observed at each station from day to day during the month of January 1957—the curves have been smoothed by taking 3-day running means. It will be seen that at five of the eight stations (namely Florida, Stanford, Sydney, Mt. Gambier, and Carnarvon) three maxima occurred on approximately the same dates in each locality; at a sixth (Perth) additional peaks were recorded, but the first and last were in general agreement with those at the above five stations. Two stations, Caloundra and Norfolk Island, clearly differed from the

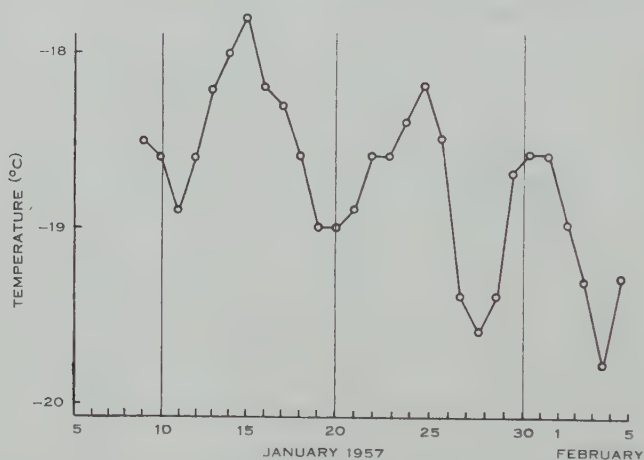


Fig. 2.—The mean temperature for 0.1 freezing nuclei per litre for all eight stations taken together.

remainder, although they agreed amongst themselves in some particulars. Both Caloundra and Norfolk Island were under the influence of a tropical cyclone in the Coral Sea during the latter part of January and it is possible that the extensive cap of ice crystals known to be generated above typical tropical cyclones may have had some bearing on the results obtained at these two stations.

The results for all eight stations have been combined to give the curve of Figure 2, which shows well-defined maxima on January 15, 25, and 31. These are consistently later than those observed in previous years by 2, 3, and 1 days respectively.

The results for the six Australian stations taken together give the mean curve of Figure 3 (b), while those for the two American stations give the curve of Figure 3 (a). It is apparent that the peak values tend to occur on approximately the same dates in the two different hemispheres.

On the basis of the 1957 measurements it may therefore be concluded that, although variations certainly exist from one locality to another, in the mean the results conform to expectations in that they exhibit three maxima in the latter

part of January and show a similarity as between the northern and southern hemispheres. The actual dates of the maxima differ from the predicted dates by 2, 3, and 1 days respectively. This may be interpreted either as being contrary to expectations or as being due to the normal fluctuations to be expected in the results for any one year. It is hoped that future measurements will show which of these two views is correct.

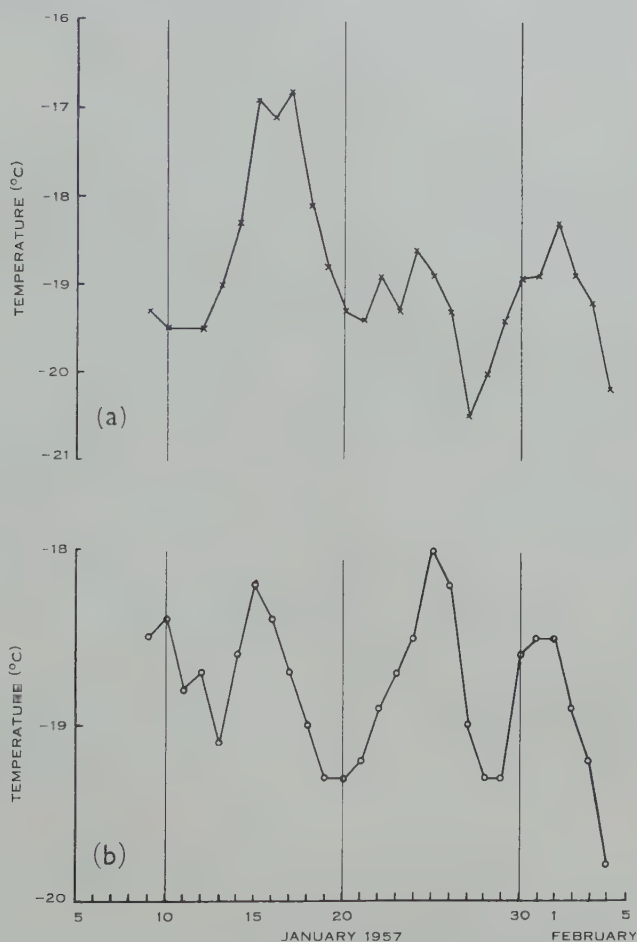


Fig. 3.—The mean curves (a) for the two northern hemisphere stations and (b) for the six southern hemisphere stations.

The author is very much indebted to Dr. R. N. Bracewell for permission to use the Stanford measurements and to the U.S. Weather Bureau for the Florida data.

References

- BOWEN, E. G. (1956).—*Nature* **177**: 1121.
 WARNER, J. (1957).—*Bull. de l'Observatoire du Puy de Dôme* No. 2: 33.

THE CONSTANCY OF THE VELOCITY OF LIGHT

By G. BUILDER*

[*Manuscript received June 2, 1958*]

Summary

The principle of the constancy of the velocity of light has generally been given a prominent place in discussions of the restricted theory of relativity. Yet lack of any clear and unambiguous statement of the principle has led to its frequently being misunderstood and to its sometimes being adduced as a basis for fallacious conclusions.

The principle is discussed from a historical point of view, is carefully analysed, and is stated in a form which seems free from ambiguity. In the context of the restricted theory the "velocity of light" must always be interpreted as "the velocity of light measured in an inertial reference system" and it must be understood that this measured value is the average speed measured over a go-and-return path. The principle can then be stated in the form: "The principle of relativity precludes any possibility of ascertaining how light is propagated relative to any inertial reference system. Measurements made in inertial reference systems, using the methods of measurement prescribed by the restricted theory, always give the same value c for the speed of light irrespective of the direction of its propagation and irrespective of the motion of its source."

Fallacious inferences about the propagation of light, arising out of an incorrect definition of "relative velocity" and out of the incorrect appellation of the relativistic law of transformation of velocities as the "law of addition of velocities" are discussed; and it is shown that the ballistic theory of light is quite incompatible with the restricted theory.

Finally, it is pointed out that the measurement procedures prescribed by the restricted theory are conventional; failure to recognize this has led many exponents of the restricted theory to assert, without sufficient justification, that these procedures demand a complete revision of the older concepts of space and time.

I. INTRODUCTION

The "principle of the constancy of the velocity of light" has generally been given a prominent position in expositions of the restricted theory of relativity. Following the lead given by Einstein in his early papers (1905, 1907), it appears frequently as one of the fundamental postulates of the theory. In most other cases it is presented as a consequence of the Lorentz transformations, when these have been otherwise established.

In spite of this, it does not seem possible to derive from the literature a statement of the principle that is clear and unambiguous and that would command universal approval. Indeed, it seems apparent that it is taken by various authors to mean quite different things. The consequent vagueness as to the exact significance of the principle leaves it open to quite improper applications. This is so not only for laymen and philosophers but even for scientists themselves.

* School of Physics, University of Sydney.

The present investigation shows that misapplication of the principle has led to some quite erroneous inferences. It will be seen that the inherent difficulty that has to be dealt with lies in the translation of the precise mathematical statement of the principle, taken together with its proper context, into a verbal statement which is not open to misinterpretation. The difficulty is therefore similar to, but far more complex than, that which led to the occurrence of the "clock paradox" of the restricted theory of relativity, and which was resolved by showing that the mathematical definition of the "rate" of a moving clock had been translated into a verbal statement lacking the precision necessary to prevent its being misapplied (Builder 1957).

To avoid any confusion in the ensuing discussion, it is necessary to state precisely what we will mean by the *restricted theory of relativity*. There is only one statement of this theory that can command universal assent, that is subject to experimental verification, and that is equally appropriate whether one ascribes the theory to Poincaré and Lorentz or to Einstein, namely,

"The restricted theory of relativity is the theory that the spatial and temporal coordinates of events, measured in any one inertial reference system, are related to the spatial and temporal coordinates of the same events, as measured in any other inertial reference system, by the Lorentz transformations."

Every prediction of the theory, thus defined, that has yet been tested, has been verified.

Thus defined, the theory is a verifiable statement about measurements made in inertial reference systems. It does not offer any causal explanation for the validity of the Lorentz transformations as relations between measurements made in different systems; it is derivable, as Einstein showed, from two postulates which are verifiable statements about the characteristics of natural phenomena, and the validity of this derivation is completely independent of any hypotheses or theories which might be held to give a causal explanation of these characteristics.

The theory was developed independently by Poincaré and Lorentz (Whittaker 1953) and by Einstein (1905). The points of view adopted in these two lines of development were so markedly different* that it will be necessary here to consider both.

II. SILBERSTEIN'S STATEMENT OF THE PRINCIPLE

To illustrate the nature of our problem we take, as an example, the form of statement of the principle given by Silberstein (1922) and typical of a number of serious and responsible authors.

* This applies more particularly to Einstein's early papers. His views tended to approach those of Poincaré and Lorentz somewhat more closely with the passage of years (Builder 1958); but we are here more concerned with his early views because these have been so widely quoted and adopted by exponents of relativity theory that they may be said to have dominated the now-current attitudes to the subject.

Silberstein's Statement

"Light is propagated *in vacuo*, relatively to any inertial reference system, with a velocity c , constant and equal for all directions, no matter whether the source emitting it is fixed or moving with respect to that system."

He goes on to say: "This is shortly referred to as uniform and isotropic light propagation in any inertial system."

To appreciate the spectacular character of this statement, it is necessary to recall that the Maxwell-Lorentz theory supposed that light is propagated in a universal stationary medium (ether) with a velocity c characteristic of the medium and independent of the motion of its source. On this theory it would seem to be incontrovertible that the propagation of the light, relative to any inertial system in motion relative to the ether, would be non-uniform and anisotropic, i.e. it would be propagated relative to the inertial system with different speeds in different directions. This inference fits in with our commonsense view of motion.

Silberstein's statement explicitly denies this anisotropy of light propagation relative to any inertial system.

Is this what Silberstein meant to convey by his statement? This question can be answered definitely in the affirmative by considering how he applied the principle in discussing the ether hypothesis. His argument (Silberstein 1924) can be set out in the form:

- (a) The restricted theory of relativity requires acceptance of the principle of the constancy of the velocity of light.
- (b) According to the Maxwell-Lorentz theory it is necessary to postulate the existence of a universal, all-pervading, stationary ether to account for the phenomena of electrodynamics and, in particular, to account for the fact that the propagation of light is independent of the motion of its source.
- (c) Yet according to the principle, as stated above, light is propagated isotropically with the same constant speed c , irrespective of the motion of its source, relative to every inertial reference system.
- (d) Hence, if it is necessary in the Maxwell-Lorentz theory to postulate the existence of an ether, it is equally necessary to postulate a similar ether associated with every one of the infinity of possible inertial reference systems.
- (e) This is clearly absurd and inconceivable, so that the ether concept breaks down and must be discarded completely.

This argument can be briefly summarized by the following quotation from Sommerfeld (1952, p. 235):

"In the earlier but long since discarded theory of the universal *ether*, the independence of the light wave from the state of motion of the emitting body was readily understood. . . . The constancy of velocity of light is today the only valid *remnant of the ether concept*. If at present we should speak of an ether, we would have to assign a separate ether to every frame of reference, i.e. speak e.g. of a primed and an unprimed ether. We now regard Lenard's 'absolute ether' merely as a freak. . . ."

The same argument seems clearly to be implied by the numerous authors who have stated that the Michelson-Morley experiment proved the ether hypothesis to be wrong; it is difficult to envisage any other argument that would lead to this conclusion.

These arguments demonstrate clearly that Silberstein's statement of the principle was understood by him, and by others, to be a substantial statement about the propagation of light relative to each and every inertial system *in exactly the same sense* that the Maxwell-Lorentz theory claimed to be a substantial statement about the propagation of light relative to one particular inertial system at rest in a universal stationary ether. If this were not so, step (d) of the foregoing argument would fail.

Silberstein's statement, thus interpreted, is indeed revolutionary in its implications. It confounds all our commonsense and physical notions and denies to us any comprehensible physical picture of the nature and behaviour of light.

Must we then accept it as a correct and necessary consequence of relativity theory? Before doing so we must at least examine it with great care.

III. LOGICAL OBJECTIONS TO SILBERSTEIN'S STATEMENT

There are serious logical objections to Silberstein's statement if it is taken in the literal sense indicated by his own and other applications of it in discussion of the ether hypothesis. In this sense it is a statement about the propagation of light.

The principle of the constancy of the velocity of light can only be justified by means of the restricted theory of relativity. Such justification might be achieved in one of two ways. On the one hand, the *principle* may be regarded as an essential postulate of the theory and the success of the theory may be regarded as evidence for the validity of this postulate. On the other hand, the *principle* may be regarded as an inference from, and a consequence of, the restricted theory. In either case, the justification rests on the validity of the restricted theory.

Now it has been pointed out, in Section I above, that the only statement of the restricted theory that can command universal assent, and that is subject to direct experimental verification, is a statement about measurements made in inertial reference systems.

Thus, so far as it is based on the validity of the restricted theory, the principle can only be a statement about measurements made in inertial reference systems.

Any further inference about the propagation of light must therefore necessarily depend on the introduction of some additional theory or hypothesis which satisfactorily relates such measurements to the physical phenomenon of light propagation. Thus, to ascertain whether Silberstein's statement about the propagation of light is justifiable, it is necessary first to analyse the measurement techniques and procedures themselves and then to investigate how the measurement results so obtained can be used to provide information about the propagation of light.

A detailed analysis of this sort is presented in the following sections of the present paper and it is shown that Silberstein's statement, taken literally, cannot be justified. On the contrary, it is shown that the *principle*, as a statement about measurements based on the restricted theory, taken together with the only available physical theory of the propagation of light, i.e. the Maxwell-Lorentz theory, leads to quite different conclusions.

It is also worth noting that, in any case, it is fairly obvious that Silberstein's statement could not be a substantial statement about the behaviour of light in *exactly the same sense* as would the corresponding statement of the Maxwell-Lorentz theory. This theory claimed to give a description of the behaviour of light about which all observers would agree, after having made appropriate measurements and after having made appropriate corrections for the effects on their measurements of their own motions. In this, the Maxwell-Lorentz theory itself provided the necessary link between the measurements and the phenomena under investigation.

The character of Silberstein's statement is quite different. It asserts that light is propagated isotropically with speed c relative to any and every inertial reference system. But this is not a statement that can command universal assent, for the following reasons.

Consider any arbitrarily selected inertial reference system S . According to the restricted theory, measurements made by observers in S of the speed of light will all give the value c , irrespective of the direction of propagation of the light and irrespective of the motion of its source. These measurement results would therefore be compatible with the S -observers supposing that light is propagated isotropically relative to their system; but they would not be proof of such isotropy without an additional hypothesis according to which the measurement results are direct evidence of physical isotropy.

Consider now a second inertial reference system S' in motion relative to S with speed v . Measurements by observers in S' will also give always the value c for the velocity of light relative to their system, and would be compatible with their also supposing the light to be propagated isotropically in their system. However, their measurements of the velocity of light relative to the system S would give values ranging between the limits $c+v$ and $c-v$, depending on the direction of the light propagation relative to the direction of the observed motion of S . These measurements would be compatible with the anisotropy of propagation of the light relative to S .

Now, if one attempted to persuade the S' -observers that the light is in fact propagated isotropically relative to S , with the constant velocity c , they could reconcile this with their own measurements only by making an appropriate correction for their own motion relative to S . This correction would, however, result also in their corrected measurements giving different values for the velocity of light in different directions relative to their own system.

It is therefore difficult to see any way in which the S' -observers could be persuaded that their measurements are compatible with the same rays of light being propagated isotropically relative both to their own system S' and to the

system S . Thus predictions of the restricted theory can only show that the measurements made in each inertial reference system of the velocity of light are compatible with isotropic propagation of light relative to that system and anisotropic propagation of light relative to every other such system.

The only possible objection to the foregoing argument is that it is at fault in stating that, according to the measurements of the S' -observers, the speed of light relative to the system S will vary between the limits $c+v$ and $c-v$, depending on the direction of propagation of the light. The basis for such an objection would be that, in calculating the speed of the light relative to S , the S' observers should have used the so-called "relativistic law of addition of velocities". It is, however, shown in Section VII below that this objection must be rejected because the application of the so-called law in such a calculation is not permissible.

It is therefore concluded that Silberstein's statement cannot command universal assent and cannot therefore be a substantial statement about the physical behaviour of light in exactly the same sense as the corresponding statement of the Maxwell-Lorentz theory. Thus step (d) in Silberstein's argument, as set out in Section II above, cannot be sustained.

IV. POINCARÉ AND LORENTZ

The Maxwell-Lorentz theory envisaged, and indeed required,* the existence of a universal stationary ether as the bearer of electromagnetic fields and as the medium of propagation of disturbances of such fields with a definite velocity c characteristic of the ether and independent of the motion of the source of the disturbance.

The principle of the constancy of the velocity of light could therefore be stated in terms of Maxwell-Lorentz theory in the form:

Statement according to the Maxwell-Lorentz theory

According to the Maxwell-Lorentz theory light is propagated isotropically in a universal homogeneous isotropic and stationary medium in a manner uniquely determined by the properties of the medium, and therefore independent of the direction of propagation and of the motion of the source of the light.

This is obviously a substantial statement about the physical characteristics and nature of light. Thus it was to be expected that an observer in uniform motion relative to the ether would be able to detect his motion by detecting the anisotropy of the propagation of light relative to him, even if the light source were moving with him.

* The Maxwell-Lorentz theory, as such, is stated explicitly in terms of such an ether. The equations of this theory, in which the velocities are defined as velocities relative to the ether, must therefore be distinguished sharply from the relativistic equations of electrodynamics, which are identical in form, but in which the velocities are defined as velocities measured in the particular inertial reference system being used. The corresponding distinction between the Fitzgerald-Lorentz contraction, defined as a contraction caused by motion relative to the ether, and the observable relativistic length contraction, is also desirable. Whether the ether hypothesis remains essential in a causal relativistic description of electrodynamical phenomena is a wider question with which I have dealt elsewhere (Builder 1958); but whatever the final answer to this question, these distinctions remain desirable.

This expectation was, of course, refuted by the Michelson-Morley experiment and by all similar experiments which followed it, thus preserving the principle of relativity of uniform motions which Poincaré (1904) restated in the form :

“The principle of relativity according to which the laws of physical phenomena should be the same, whether for an observer fixed, or carried along in a uniform motion of translation, so that we have not and could not have any means of discerning whether or not we are carried along in such a motion.”

For brevity this will here be referred to as the *principle of relativity* since the present context naturally excludes any wider connotation of this term.

The task faced by Poincaré and Lorentz was therefore the reconciliation of :

- I. The principle of relativity.
- II. The principle of the constancy of the velocity of light, according to the Maxwell-Lorentz theory, as stated above.

These may therefore be regarded as the postulates from which they derived the restricted theory.

Neither postulate could very well be discarded. The experimental evidence in favour of the principle of relativity was overwhelming. On the other hand, the Maxwell-Lorentz theory had, in all other respects, proved to be a wholly successful description of all the available empirical data relating to light and electrodynamics. There was no available alternative. The ballistic theory of Ritz (1908) is incompatible with these data ; in particular it is incompatible with the fact, now securely established experimentally, that the velocity of light is independent of the motion of its source. This last is further discussed in Section VIII below.

Poincaré and Lorentz, by work extending over the period 1892–1904, succeeded in establishing the restricted theory of relativity and thus reconciling the two postulates given above ; a detailed account of this development is given by Whittaker (1953) and need not detain us here. In achieving this they did not need to modify the concept of the ether in the Maxwell-Lorentz theory nor did they need to modify the principle of the constancy of the velocity of light according to that theory. The ether hypothesis was not only retained ; it also provided the basis for a causal explanation for the fact that measurements made in different inertial reference systems are related by the Lorentz transformations.

The hypothesis, put forward independently by Lorentz (1892) and by Fitzgerald (see Lodge 1893), that bodies in motion relative to the ether with speed v are contracted, in the direction of their motion, by the factor $\sqrt{1-v^2/c^2}$, was sufficient by itself to account for the negative result of the Michelson-Morley experiment. Moreover, it entailed* the consequence that clocks in motion

* I have pointed out elsewhere (Builder 1958) that the Fitzgerald-Lorentz contraction and the clock-rate reduction are not two independent hypotheses. This seems first to have been shown by Larmor (1900) and can readily be illustrated by considering a simple clock consisting of a rigid rod fitted with reflectors at each end so that a ray of light will be propagated backwards and forwards along the length of the rod. It can be shown that, if the rod suffers the Fitzgerald-Lorentz contraction when moving, the frequency with which the light traverses the go-and-return path along the rod will be reduced by the factor $\sqrt{1-v^2/c^2}$.

relative to the ether with speed v should also suffer a reduction in rate by the factor $\sqrt{1-v^2/c^2}$.

The Fitzgerald-Lorentz contraction and the entailed clock-rate reduction, taken together, are sufficient to explain the fact that measurements made in different inertial reference systems can be related by the Lorentz transformations; it was only necessary to recognize also that the "local time" of Lorentz, which he had originally devised as a mathematical trick to achieve covariance of the Maxwell-Lorentz form of equations to the Lorentz transformations, is the only "time" that can be established in an inertial reference system moving with unknown velocity through the ether. Poincaré recognized this in 1904* and in the same paper prescribed the now well-known relativistic method of synchronization of clocks in any inertial reference system.

The Poincaré-Lorentz postulates I and II themselves *entail the impossibility of ever discovering how light is propagated relative to any inertial reference system*. In other words, *they preclude the possibility of making any measurements that could reveal whether or not light is propagated isotropically relative to any particular inertial reference system, or of revealing the degree of anisotropy*. This is obvious: propagation of the light in the ether is isotropic (postulate II) but the detection of anisotropy in the propagation relative to an inertial reference system would reveal the motion of the system relative to the ether in contravention of the principle of relativity (postulate I).

More specifically, this entails the impossibility of measuring the velocity of light relative to any inertial reference system by determining the time taken for a light signal to travel over a unidirectional path from one point in the system to another, i.e. the impossibility of measuring the unidirectional velocity of light over a one-way path in the system. For were such measurements possible they would reveal any anisotropy of light propagation relative to the system in contravention of the principle of relativity.

Let us suppose that observers in an inertial reference system were to attempt such measurements. They would have to establish the facilities necessary to measure the time taken for a light signal to travel over a measured distance from one point A in the system to another point B . This would require having at A and B clocks known to be synchronous.† To achieve such synchronization it would be necessary to relate the readings of the clock at A to the readings of the clock at B by some signalling method. The fastest available signal is a

* Minkowski (1908) was incorrect in ascribing the first recognition of this to Einstein. On the other hand Einstein (1907) stated, as quoted in Section VI below, that this recognition was all that was essential to solve the basic problem.

† It is sometimes erroneously stated that measurements of light velocity over a unidirectional path can be made in a terrestrial laboratory without this provision, e.g. by the use of a pair of toothed wheels running synchronously on a common shaft. This is fallacious. It presupposes that the common shaft is a guarantee of synchronization. This is not so. The shaft cannot be set in rotation by torques applied synchronously at its two ends without having available two synchronized clocks to ensure that the application of the torques is synchronous. If it is set in rotation by a torque at one point, the time required for transmission of the torque along the rod will upset the synchronization by an amount which could not be ascertained unless synchronized clocks were available. This remains true however short the shaft may be.

flash of light; thus it would be necessary to allow for the time of transmission of the light signal from point *A* to point *B*, or vice versa. To make this allowance it would be necessary to know the velocity of light, in each direction, relative to the system. *Thus the clocks cannot be synchronized unless the unidirectional velocity of light relative to the system is known, and the unidirectional velocity of light relative to the system cannot be measured without synchronizing the clocks.* It can be shown that all other methods of synchronization, e.g. by slow transport of clocks from one place to another, are subject to the same limitations.

It clearly follows that when we speak, as we often do, of the "velocity of light measured in an inertial reference system" we cannot be speaking of a measurement of the unidirectional velocity; nor can we be speaking of a measurement which contains any information about the propagation of light relative to the system or which can be used as a basis for any substantial statement about the propagation of light relative to the system.

What then are we to understand by "the velocity of light measured in an inertial reference system"? *It means the average velocity of a light signal propagated over a go-and-return path.*

This average velocity is obviously measurable. This requires only the use of a single clock, located at the point *A* of emission of a light signal, to measure the time taken for the signal to reach a distant point *B* in the system and to return to the point *A* after reflection at *B*. Knowing the time taken and the total distance travelled, the average velocity can be calculated immediately.

Since this average velocity is measurable in any inertial reference system, the principle of relativity (postulate I) requires that its measured value must be the same for all directions in any one system and must have the same value c in all such systems as in the ether. Otherwise such measurements, e.g. using the Michelson-Morley experiment, would enable the motion of the system, relative to the ether, to be detected, in contravention of the principle of relativity.

This is the principle of the constancy of velocity of light of the restricted theory. To avoid confusion it should be referred to more explicitly as *the principle of the constancy of the measured average value of the velocity of light over go-and-return paths*, and it should be stated in some such form as the following:

The principle of relativity precludes any possibility of ascertaining how light is propagated relative to any inertial reference system. The only measurements possible are measurements of the average speed of light over a go-and-return path and these always give the same value c irrespective of the direction of transmission of the light and irrespective of the motion of its source.

To this it may perhaps be objected that observers in an inertial reference system can in fact make measurements of the speed of light propagated over a one-way path in their system when once they have synchronized their clocks by the method specified by Poincaré (1904) and by Einstein (1905) and that, if they do so, they must, in accordance with the principle of relativity, always obtain the value c .

This objection cannot be sustained. Although, in such a measurement procedure, the light is propagated over a one-way path from one point *A* in the

system to another point B , the result is still a measurement of the average velocity over a go-and-return path from A to B because of the procedure, necessarily involved in the measurements, of synchronizing the clocks at A and B .

The prescribed method of synchronizing the two clocks is as follows. Let a ray of light be emitted from A when the clock there reads t_A ; let it reach B when the clock there reads t_B and let it be then reflected back to reach A again when the clock there reads t'_A . The two clocks are synchronized if

$$t_B - t_A = t'_A - t_B.$$

In other words, the clocks are to be set so that measurements made by these clocks will give the same time for transmission from A to B as from B to A . Thus, according to these clocks, the time taken for light to travel in either direction will be equal to the time it would take if it were propagated at a speed equal to its average speed from A to B and from B to A .

This method of setting the clocks is a *convention* which leads to a *conventional measure* of the velocity of light from A to B or from B to A . This conventional aspect of the restricted theory is discussed in Section IX below. Thus, in the context of the restricted theory, the *velocity of light* must be interpreted as a conventional reference to the measured average value of the speed of the light over a go-and-return path.

In this context, measurements of the "velocity of light" clearly lack any simple and direct relation to the physical propagation of light relative to the inertial reference system in which the measurements are made. If we do wish to infer from such measurements anything about the propagation of light relative to an inertial reference system, the only tenable physical theory at our disposal is that of Maxwell and Lorentz. And it has been shown above that the results of such measurements are predicted by, and are therefore compatible with, this theory.

V. EINSTEIN (1905)

It can readily be shown that the conclusions reached in the last section are completely consistent with the expositions of the restricted theory given by Einstein (1905, 1907).

In Section 1 of his 1905 paper, Einstein sets out the following statements about the synchronization of two clocks A and B at rest at different points in an inertial reference system :

- (i) "A common time for A and B . . . cannot be defined at all unless we establish *by definition** that the 'time' required for light to travel from A to B is equal to the 'time' it requires to travel from B to A ."
- (ii) "Let a ray of light start at the ' A -time' t_A from A towards B , let it at the ' B -time' t_B be reflected at B in the direction of A , and arrive again at A at the ' A -time' t'_A ."

In accordance with definition the two clocks synchronize if

$$t_B - t_A = t'_A - t_B."$$

* The italics are Einstein's.

(iii) "In agreement with experience we further assume the quantity

$$\frac{2AB}{t'_A - t_A} = c$$

to be a universal constant—the velocity of light in free space."

His first statement (i) expresses the conclusion reached in Section IV that clocks can be synchronized by light signals in each inertial system only if we accept a conventional *definition*, that the "time" required for light to travel from *A* to *B* is the same as that required for it to travel from *B* to *A*. This is equivalent to Poincaré's assertion (1904) that the "local time" of Lorentz is the only time that can be established in an inertial reference system and, like it, is convenient because the principle of relativity precludes our ever discovering how light is propagated relative to the system.

His second statement (ii) prescribes the method of synchronization, previously given by Poincaré in 1904, for establishing the "local time" of Lorentz in the system.

*His third statement (iii) declares explicitly that the quantity c is equal to the average velocity of light over the go-and-return path from *A* to *B* and back to *A*, i.e. it is twice the distance AB divided by the time $t'_A - t_A$, measured on a single clock, for the go-and-return transmission. Moreover, it declares that in agreement with experience, i.e. the results of the Michelson-Morley and other experiments, this quantity c is a universal constant. This is in precise agreement with the conclusion reached in Section IV that the principle of relativity precludes our ascertaining how light is propagated relative to an inertial reference system but requires that the average velocity of the light, measured over a go-and-return path, shall have the same value c in all such systems.*

Statement (iii) concludes with a parenthetical phrase identifying the universal constant c with the "*velocity of light in empty space*". The meaning of this is by no means clear. It might, on the one hand, be taken to be a definition of the "velocity of light in empty space" in any inertial reference system; if this is so, it is clear that this definition precludes any possibility of this "velocity" implying any information about the anisotropy of light propagation relative to such a system. It might, on the other hand, be taken as an identification of the quantity c with the velocity of light in the ether according to the Maxwell-Lorentz theory; but this interpretation seems to be excluded by Einstein's claim, in the introductory paragraphs of the same paper, that the concept of the ether is superfluous.*

Thus Einstein's statements (i), (ii), (iii) are completely compatible with, and may even be held to express concisely, the conclusions reached in Section IV above. In particular, they are clearly compatible with the statement there given of the principle of constancy of light velocity and the corresponding interpretation of the "velocity of light" in the context of the restricted theory.

* Though he did not long adhere to this view (Builder 1958), the fact that he did indeed hold it in 1905 is confirmed by his more emphatic statement in his 1907 paper.

Yet, in Section 2 of his paper, Einstein immediately, without any further definition of terms, sets out his two postulates, which he refers to as the *principle of relativity* and the *principle of the constancy of the velocity of light*, as follows :

- I. "The laws by which the states of physical systems undergo change are not affected, whether these changes of state be referred to the one or the other of two systems of coordinates in uniform translatory motion."
- II. "Any ray of light moves in the 'stationary' system of coordinates with the determined velocity c , whether the ray be emitted by a stationary or by a moving body. Hence

$$\text{velocity} = \frac{\text{light path}}{\text{time interval}}$$

where the time interval is to be taken in the sense of the definition of Section 1."

The "stationary" system he had defined in Section 1 as any arbitrarily selected inertial reference system.

Taken literally, the statement of postulate II is, like that of Silberstein, a statement about *how light moves* relatively to an arbitrarily selected inertial reference system. Taken thus, it asserts unequivocally that light *is propagated* isotropically relative to such a system.* This is undoubtedly how it has been interpreted by Silberstein and others, as discussed in Section II above.

Yet the context shows that this literal interpretation is quite untenable because it is wholly incompatible with the statements (i), (ii), (iii) quoted above from Section 1 of his paper. The possibility of making any assertion about how "light moves" in the system had been specifically denied in statement (i). The "velocity c " of light had been defined in statement (iii) as the average speed over a go-and-return path.

Thus, in the full context, we can only properly interpret postulate II as a statement in which the "*determined velocity c* " is that defined by statement (iii). We must therefore infer that the statement that "*light moves*" with the "*determined velocity c* " can be interpreted only to mean that measurements of the velocity of light rays, made by observers in the system, utilizing the conventions prescribed in statements (i) and (ii), must always give the same value c "*whether the ray be emitted by a stationary or by a moving body*".

* It is important to note that, taken literally, the postulate would be indistinguishable from the *principle of constancy of light velocity according to the Maxwell-Lorentz theory* if one supposed the "stationary" system to be at rest in the ether. The "time interval", "taken in the sense of the definition of Section 1", would then be identical with the "absolute" time of a system at rest in the ether.

Nor would Einstein's derivation of the restricted theory be in any way affected by such an identification *because the definition of the stationary system does not enter into, or affect, his utilization of the postulate*.

It is indeed difficult to avoid the feeling that Einstein's use of "stationary" was an unconscious reflection of the fixed ether of the Maxwell-Lorentz theory. That this may have been the case is to some extent supported by the introductory paragraphs of his 1907 paper.

Thus interpreted, postulate II becomes, once again, a statement about measurements, and not a statement about the propagation of light relative to the system. And it expresses the conclusion, required by the principle of relativity of postulate I, that such measurements must fail to reveal any anisotropy of light propagation relative to the system. Once again, any inference from the postulate about the propagation of light relative to the system would require a further physical theory about the relation of the measurements to the phenomenon of propagation. Such a theory would lie outside the context of the restricted theory as it was defined in Section I; in any case, the only tenable theory available is that of Maxwell and Lorentz.

VI. EINSTEIN (1907)

The conclusions reached in the last two sections are further supported by reference to Einstein's second comprehensive paper on the restricted theory, published in 1907.

In the introduction to this paper he again specifically rejected the ether hypothesis. Moreover, he rejected the Fitzgerald-Lorentz contraction hypothesis as an *ad hoc* assumption and an artificial device to rescue the Maxwell-Lorentz theory from the results of the Michelson-Morley experiment. He then states:

"However, it turned out surprisingly that it was only necessary to define sufficiently precisely the concept of time to overcome these difficulties. It required only the recognition that the auxiliary 'local time' of Lorentz can be defined simply as the 'time'. . . The Fitzgerald-Lorentz hypothesis then appears as a necessary consequence of the theory."

I have quoted this statement in full partly because of its importance in the assessment of the historical significance of Poincaré's recognition of this in 1904, as discussed in Section IV above.

It is, however, also important as contextual background for his subsequent treatment of the problem of clock synchronization. He assumes an inertial reference system to be equipped with ideal standard clocks and, without further discussion, simply states:

"We now assume that the clocks can be so adjusted that the velocity of propagation of a light ray in vacuum—measured with the aid of these clocks—will everywhere be equal to a universal constant c .

If A and B are two points fixed in the coordinate system . . ., whose separation is r , and if t_A is the reading of clock A when a light ray is emitted in the direction AB , and t_B the reading of the clock B on arrival of the light signal then, irrespective of the motion of the light source,

$$r/(t_B - t_A) = c.$$

That the assumption here made, which we will call the 'principle of the constancy of the velocity of light' should be satisfied in nature is by no means obvious, yet this . . . is made plausible by the confirmation it has been given by experiments."

This restatement by Einstein confirms entirely the interpretation, given in Section V above, of his 1905 paper.

The principle, as he now states it, is simply "*that the velocity of propagation of a light ray—measured with the aid of these clocks—will everywhere be a universal constant c* " and it is here clearly given as a *prescription of how the clocks shall be set*.

It has thus clearly become a statement about a conventional procedure of measurement. Any simple and direct inferences about the propagation of light relative to inertial reference systems have been omitted and have, in fact, been clearly precluded by the context.

VII. THE RELATIVISTIC LAW OF TRANSFORMATION OF VELOCITIES

The principle of the constancy of the velocity of light can be inferred from the relativistic law of transformation of velocities. This law is itself simply derivable from the Lorentz transformations and provides a relation between the velocity of a thing as measured in any one inertial reference system S and the velocity of the same thing as measured in any other inertial reference system S' . In particular, if the measured value of the velocity of light in the system S is c , the transformation gives the same value c for its velocity measured in the system S' . Thus the principle of the constancy of the velocity of light appears once again as a statement about measurements made in inertial reference systems, in agreement with the conclusions reached in previous sections.

Nevertheless, it is necessary to discuss the law in some detail here because it has sometimes been misapplied and has thereby led to some erroneous inferences which might be presented as objections to the conclusions reached in previous sections. Such inferences have arisen out of a serious ambiguity in the meaning of the term *relative velocity*. It is therefore necessary to set out here a definition of this term (to be referred to as the *old* definition) and to show that an alternative definition sometimes used (the *new* definition) is untenable.

The velocity of a thing is its velocity stated in terms of the coordinates of some specified inertial reference system S . We will refer to it as *the velocity of the thing measured in S* or, more briefly, as *the velocity of the thing in S* . This definition is not to be taken to imply the existence of any physical system corresponding to the inertial reference system S . On the contrary, the system S is to be understood primarily as a system of coordinates in the mathematical sense, and the phrase "measured in S " is to be understood primarily to mean "expressed in terms of the measures, i.e. coordinates, of S ". It is true that the actual measurement of such a velocity implies the use of some physical system corresponding to at least one inertial reference system but, once this measurement has been made, the velocity in any other inertial reference system may be calculated by utilizing the relativistic law of transformation.

The relative velocity of two things is the velocity of one thing A relative to another thing B , measured in some specified inertial reference system S as the simple vector difference $\mathbf{u}_a - \mathbf{u}_b$ between the velocities \mathbf{u}_a and \mathbf{u}_b of the two things in S (*old definition*).

Thus a relative velocity measured in S is a simple vector relation between the velocities in S of two real things such as bodies, light quanta, light rays, etc., and is to be calculated by the parallelogram or vector law of addition. Thus any

statement that implies such a relation is essentially a statement about relative velocities. To illustrate this, consider the postulate that "the velocity of light is independent of the motion of its source". This is a statement of relation between the velocity of light and the velocity of its source, both measured in the one inertial reference system. It could obviously be restated, explicitly in terms of relative velocities, in the clumsy but equivalent form: "If a ray of light is propagated in any direction \hat{i} in S , after being emitted from a source moving with velocity \mathbf{v} in S , its velocity relative to its source, as measured in S , has the value $c\hat{i} - \mathbf{v}$, whatever the value of \mathbf{v} and whatever the direction \hat{i} ."

Now, if one of the things B , referred to in the definition, is at rest in S , so that $\mathbf{u}_b = 0$, the velocity of A relative to B , measured in S , is $\mathbf{u}_a - \mathbf{u}_b = \mathbf{u}_a$ and is equal to the velocity \mathbf{u}_a of A in S . It would therefore seem that one could, if one wished, speak of the velocity \mathbf{u}_a of A in S as *the velocity of A relative to S measured in S* , though it would be a rather superfluous and rather clumsy way of referring to what is already sufficiently and concisely described as *the velocity of A in S* . It might, however, properly be objected that the concept of relative velocity is essentially one of relation between the motions of two real things, whereas the inertial reference system S does not necessarily imply the existence of any physical system or of any real thing at rest in S .

However this may be, it will be shown that *it is certainly not permissible* to refer to the velocity of A in S as *the velocity of A relative to S* or as *the velocity of A relative to some thing at rest in S* . To do so would be to utilize a *new* definition of relative velocity which will be shown to be untenable. This *new definition* would state that *the velocity of A relative to B is the velocity of A in the rest system of B , i.e. expressed in terms of the measures of the inertial reference system in which B is, at least momentarily, at rest*. It is to be noted that this would exclude from the concept of relative velocity the corresponding relation between the velocities of A and B measured in any other inertial reference system. In other words, it would imply that the only significant relation between the velocities of bodies is the relation between their velocities in the rest system of one of them.

The consequences of using this *new* definition can be illustrated thus. According to it, the velocity of light relative to its source is its velocity measured in the rest system of the source. The velocity of light is c in every inertial reference system. Therefore the velocity of light relative to its source is c , and this would remain so irrespective of changes in the motion of the source. We would then have, as a *consequence* of the restricted theory, the statement that "the velocity of light relative to its source is c irrespective of the motion of the source", while we have, as a *postulate* of the theory, the statement that "the velocity of light is independent of the motion of its source", and we have shown above that this can only mean that the velocity of light relative to its source depends on the motion of its source.

This apparent contradiction is of course due to the fact that different definitions of relative velocity have been used in the two statements. To avoid such contradictions we must make a choice between these definitions. *There are grave objections to the new definition.*

In the first place, it should be noted that in the context of Newtonian relativity the *new* and *old* definitions would always result in the same value for the relative velocity, simply because the measured value of the relative velocity would be the same for all inertial reference systems ; it might therefore be thought that there would be no essential difference between the two definitions. But this is not so. For, even in this context, the *new* definition conceptually contravenes a principle which is basic in the formulation of the laws and equations of physics, i.e. *that in every statement of a physical law, in every physical equation, and in every description of natural phenomena, all the quantities referred to must be stated in terms of the measures of one and the same reference system.* The necessity for this principle was in no way affected by the fact that the laws and equations of Newtonian mechanics retained the same form in every inertial reference system ; indeed this covariance itself depends on all the quantities specified in the laws and equations being measured in the particular reference system being used. Thus the new definition is conceptually inadmissible in the context of Newtonian relativity even though it would not lead to obvious contradictions.

But in the context of the restricted theory the relations between velocities are no longer independent of the reference system. Thus any departure from the principle of stating these relations in terms of the measures of the single reference system implied in any physical statement, whether it be a law or an equation or a description of phenomena, can only lead to chaos. The *new* definition clearly contravenes this principle in the worst possible way in that it leads to physical statements in terms of the measures of an unspecified multiplicity of unspecified inertial reference systems, e.g. to a statement about the velocity of light relative to its source, measured in all the rest systems of a source of which the motion is completely unrestricted and unspecified.

In the second place, it is to be noted that the rejection of the *old* definition and the adoption of the *new* would be abortive. To illustrate this, consider once again the postulate that "the velocity of light is independent of the motion of its source". This is a description, in terms of the measures of any one arbitrarily selected reference system, of a physical relation which could be expressed by saying that "the motion of light is independent of the motion of its source". If we were to adopt the *new* definition and reject the old, we would then have to formulate a description, in terms of the *new* definition, corresponding to this physical relation. Since the only statement of relation that is admissible according to the *new* definition is the relation measured in the rest system of the source, we would be forced to the statement that "the velocity of light relative to its source always has the same value c , irrespective of the motion of its source" ; but this statement fails to characterize uniquely the postulated physical relation, for it would remain true even if the emission were ballistic, i.e. if the motion of the light were determined by the motion of its source.

We thus conclude that the *new* definition must be rejected and the *old* definition retained in the context of the restricted theory.

Thus we must continue to calculate relative velocities by the simple vector law, i.e. the parallelogram law, of addition of velocities, even though we must replace this law by the relativistic law of transformation for the transformation

of velocities from the measures of one inertial reference system to the measures of another.

Thus the claim, frequently made, that the parallelogram law of addition of velocities has been replaced by the relativistic law of transformation is true *only* in regard to transformations of velocities from the measures of one inertial reference system to the measures of another. It is not true in regard to calculations of relative velocities. It is perhaps worth also making the obvious, but not trivial, remark that it is certainly not true of the resolution and composition of velocities which is the basis of all mathematical formulations of kinematics; in spite of the restricted theory we continue freely and without embarrassment to use the parallelogram law to resolve velocities into their tangential and normal components, to resolve them into their Cartesian components, to add such components, or to add the velocities of two or more simple harmonic motions, and so on.

Thus the practice of referring to the relativistic law of transformation of velocities as the "*relativistic law of addition (or composition) of velocities*", although originated by Einstein (1905, Section 5, "The Composition of Velocities"; 1907), and, although adopted by many authors such as Whittaker (1953) and Sommerfeld (e.g. 1952), is a misnomer which is both pointless and dangerously misleading. It is pointless because it is, at best, an inaccurate description of what is in fact a law of transformation. It is dangerous because it may be taken to imply (as it was meant to imply) that the law is to be used for purposes other than transformation. The *new* definition of relative velocity, discussed above, was also suggested in Section 5 of Einstein's 1905 paper when he, apparently quite casually, referred in effect to *velocity in S* as *the velocity relative to S*.

The appalling confusion to which this has led can be well illustrated by a further example. Eddington, who was precise and careful in his formal presentations of relativity theory, made the following statement in a popular exposition (1928).

"A feature of the relativity theory which seems to have aroused special interest among philosophers is the absoluteness of the velocity of light. In general velocity is relative . . . But it is a curious fact that if I speak of a velocity of 299,796 kilometres per second it is unnecessary to add the explanatory phrase. Relative to what? Relative to any and every star or particle of matter in the universe."

It is in fact necessary to add a number of explanatory phrases, i.e. to avoid this statement being taken seriously by laymen and philosophers or even by physicists.

The fallacy is obvious. The argument leading to the statement must take the form: The old parallelogram law has been replaced by the new relativistic law of addition of velocities. Therefore, to calculate the velocity of light relative to any star or relative to any particle of matter, we must use the relativistic law of addition, i.e. we must calculate the relative velocity as measured in the momentary rest system of the star or particle. When we do this we always obtain the value c irrespective of how the stars or particles are moving and irrespective of how their motions are changing.

The statement must therefore be rejected as being based on an untenable definition of relative velocity. If we translate it into proper physical terms it becomes trivial and uninteresting, for all it amounts to is that, if we were to measure the velocity of light in the rest system of a star or particle, we would obtain the value c , just as we would in any other inertial reference system.

The discussion in this section also demonstrates the fallacy in the possible objection mentioned in the penultimate paragraph of Section III.

VIII. THE BALLISTIC THEORY OF LIGHT

We have now to consider a serious misapplication of the relativistic law of transformation of velocities which, in some recent discussions of the ballistic theory of light, has led to an absurd conclusion.

In the ballistic theory it is postulated that light is emitted with a definite velocity relative to its source. Whittaker (1953, p. 38) treats the corpuscular theory as being synonymous with the ballistic theory. He states that, according to the corpuscular theory,

“the corpuscles emitted by a moving star would have a velocity which is compounded of the velocity of the star and the velocity of the light relative to a source at rest, just as an object thrown from a carriage window in a moving railway train has a velocity which is obtained by compounding its velocity relative to the carriage with the velocity of the train (the *ballistic* theory).”

There is, of course, no *a priori* reason why light thus emitted ballistically should consist of simple corpuscles totally devoid of wave-like characteristics. An analogy, between the wave-like characteristics of electrons and the corpuscular-like characteristics of light quanta, has suggested to some authors the possibility that light quanta might be emitted ballistically; but the general acceptance of the restricted theory of relativity has made it necessary for them to consider this possibility in the context of that theory.

The great interest in the ballistic hypothesis at the beginning of this century, which led to the remarkable attempt by Ritz to develop a complete ballistic theory of electrodynamics, arose out of the fact that the ballistic hypothesis is compatible with the negative result of the Michelson-Morley experiment and would, if it were tenable, provide a complete and satisfactory explanation of that result within the context of simple Newtonian relativity.

It is in fact easy to show that, in the strict logical sense, the ballistic theory of light and the principle of relativity, taken together, are *equivalent* to simple Newtonian relativity. Taken together, they *entail* this. Moreover, simple Newtonian relativity would *entail* the principle of relativity and would also *entail* the ballistic theory of light.

On the other hand, it has frequently been shown that the principle that the velocity of light is independent of the motion of its source and the principle of relativity, taken together, are *equivalent* to the relativity* of the restricted

* I here distinguish the relativity of the restricted theory from the restricted theory of relativity, in accordance with the conclusions reached in Section IX below.

theory. Taken together they *entail* this. Moreover, the relativity of the restricted theory *entails* the principle of relativity and also *entails* the principle that the velocity of light is independent of the motion of its source.

These statements are indisputable. Since the relativity of the restricted theory is incompatible with simple Newtonian relativity, it follows necessarily that *the ballistic theory is incompatible with the relativity of the restricted theory*.

This same conclusion would also follow directly from a statement that the ballistic theory is incompatible with the principle that the velocity of light is independent of the motion of its source. This statement is certainly entailed by the definitions of the ballistic theory and of the principle. Any doubt about this could arise only out of the ambiguity in the definition of relative velocity discussed in Section VII above.

The experimental evidence against the ballistic theory is quite overwhelming. Strong experimental evidence has led to general acceptance of the restricted theory of relativity and this entails acceptance of the principle that the velocity of light is independent of the motion of its source. Experimental evidence has also rendered untenable any theory of electrodynamics based on the ballistic hypothesis. Finally, it has now been well established directly by astronomical data, presented by de Sitter, Comstock, and others, as well as by terrestrial experiments, that the velocity of light is in fact independent of the motion of its source; the wealth of such evidence has been summarized briefly by Whittaker (1953).

In spite of all this, Matthias, Whittaker, and Sommerfeld have recently claimed that a ballistic theory of light emission is reconcilable with the restricted theory of relativity.

The logical incompatibility of the ballistic hypothesis and the restricted theory, demonstrated above, is sufficient justification for rejecting this claim. Yet the matter is one of such fundamental significance that it seems necessary here to analyse the claims and demonstrate the fallacies in the arguments put forward.

Whittaker (1953), having summarized the strong experimental evidence against the ballistic hypothesis, continues as follows:

"It was now recognised that these observational findings, which in the nineteenth century might have been supposed to tell in favour of the wave theory, were actually without significance one way or the other. . . . For, according to relativity theory, even on a corpuscular hypothesis, a corpuscle which had a velocity c relative to its source would have the same velocity c relative to any observer, whether he shared in the motion of the source or not."

The fallacy in this is obvious. In the context of the restricted theory, the statement that a corpuscle has a velocity c relative to its source has no definite meaning unless the reference system, in which this is claimed to be true, is specified. If it is supposed that the instantaneous rest system of the particle is implied, the statement is insufficient to entail that the emission be ballistic, for the statement is equally true of emission in accordance with the principle that the velocity of light is independent of the motion of its source. Thus the inference from this

statement, that the emitted light has the same velocity c relative to any inertial system of reference, tells us nothing about the ballistic theory.

Furthermore, if it is held that the statement is meant to imply that the corpuscle has the velocity c relative to its source as measured in any and every inertial system, this is in direct contradiction to all the experimental evidence which shows, for example, that in the inertial reference system used by astronomers the velocity of light does not have the velocity c relative to its source but has the velocity c in that inertial system. This contradiction could be avoided only by claiming that the astronomers should calculate the velocity of the light relative to its source by the "relativistic law of addition of velocities"; but it has been shown in Section VII that this would give the velocity of the light relative to its source as measured in the rest system of the source and not as measured in the reference system being used by the astronomers.

Sommerfeld's case depends directly on such a misapplication of the law of transformation of velocities. He writes (1952),

"The fact that Newton's emission theory could, in a sense, experience a resurrection in the present theory of light quanta rests solely on the addition theorem of the theory of relativity according to which $c+v=c$ (c =velocity of light quanta, v =velocity of emitting body)."

The fallacy here has been exposed in the previous paragraph and in Section VII.

The argument of Matthias (1939), though put forward in considerable detail, is essentially the same as that of Whittaker and must be rejected for the same reasons.

Thus any claim that the ballistic theory could be reconciled with the relativity of the restricted theory is absurd and must be rejected vigorously.

IX. CONVENTIONAL ASPECTS OF THE RESTRICTED THEORY

In Sections III-V above it has been suggested that the method of synchronizing clocks specified in the restricted theory is a convention. Whether or not this is true is a question of considerable interest and may even perhaps be of fundamental importance from the point of view of critical philosophy.

The obvious objection to this suggestion is that the restricted theory does in fact lead to some remarkable conclusions about the characteristics of physical phenomena, although it is derivable from only two basic postulates. For example, it leads to the conclusion that energy and inertial mass are equivalent. How then can one regard the basic measurement procedures of the theory as conventional when the powers of physical prediction of the theory are so great?

It is first of all necessary to distinguish between the restricted theory itself, in the sense defined in Section I, and the physical theories of dynamics and electrodynamics which have been formulated in terms of the measures prescribed by the restricted theory. At the same time it must be recognized that the restricted theory itself does in fact also imply some of the physical characteristics of nature; it entails that nature is such that we cannot ever, by observations of dynamical or electrodynamical phenomena, measure absolute velocity; it entails also that the propagation of light is independent of the motion of its source.

Starting from the other end, we find that the task faced by Poincaré and Lorentz, and by Einstein, was the reconciliation of the two empirical generalizations, the principle of relativity and the principle that the velocity of light is independent of the motion of its source. The essential factor in effecting this reconciliation was the recognition that we cannot, by observations of dynamical and electrodynamical phenomena, achieve measurements corresponding directly to the concepts of absolute time and of absolute space (Einstein 1907). This recognition was, by itself, sufficient to demonstrate that the two principles were not necessarily incompatible, for any measurement of absolute uniform velocity, in contravention of the principle of relativity, would necessarily require absolute measurements of space and time.

Having, by this recognition, shown that there is no fundamental incompatibility between the two principles, it remained necessary to decide what convention should then be adopted in physical measurements. Some such convention was of course necessary once it was recognized that the unique measurements implied in the absolute concepts were impossible.

The most desirable convention seemed obviously to be that of synchronizing the clocks in each inertial reference system by means of light signals or by some other means which would permit a system time to be established uniquely without need to refer at all to any other possible reference systems. This was the convention adopted in the restricted theory. In particular, the procedure adopted was that of setting the clocks to the local time of Lorentz;* this had

* It has been previously pointed out by Ives (e.g. 1951) and by Grünbaum (1955) that the adoption of this particular procedure also involves a further conventional choice. The condition for synchronization of clocks in the restricted theory may be written, using the nomenclature of Sections IV and V above, in the form $t_B = t_A + \epsilon(t'_A - t_A)$, providing that ϵ has the value $\frac{1}{2}$.

Grünbaum has pointed out (1955) that "no fact of nature found in the objective relations of physical events precludes our choosing a value of ϵ between 0 and 1 which differs from $\frac{1}{2}$ " and, after a critical discussion of possible objections to this statement, he concludes that "the value of $\epsilon = \frac{1}{2}$ is simpler only in the *descriptive* sense of providing a *symbolically* simpler representation of these data".

Similarly, Ives (1951) wrote: "A point of great importance may here be noted. It is that we do not need to assign individual values to c_o and c_b (i.e. the velocities of light in the outward and backward directions on a moving platform), such for instance as calling them equal as is done in Einstein's arbitrary 'definition' of simultaneity. We carry these quantities as real although undetermined quantities. . . ." This view was based, not on a logical analysis of the restricted theory as such, but on a careful analysis made in the course of his own critical and independent investigation, in terms of the ether hypothesis, of measurements that are possible in inertial reference systems. His generalized transformations cover the general case in which ϵ is unspecified, and they reduce to the Lorentz transformations for $\epsilon = \frac{1}{2}$.

Ives's work is valuable, and it is important in the present context. It demonstrates that the restricted theory, as defined in Section I above, can be rigorously established on the basis of the ether hypothesis and, at the same time, draws attention strongly to the conventional character of the choice $\epsilon = \frac{1}{2}$ in the Poincaré-Lorentz development of the theory.

Ives's work resulted from his rejection of the restricted theory. It is significant that this rejection was due to the fact that he, like many others, had been misled into believing that the principle of the constancy of the velocity of light, according to Silberstein's statement and interpretation as given in Section II above, is an essential feature of the restricted theory; unlike most others, he stoutly maintained the view that this was "not merely 'un-understandable', it is *not* supported by objective matters of fact; it is untenable, and . . . unnecessary". This view is in agreement with the conclusions reached in the investigation presented here.

the inestimable advantage that the equations of physics then had precisely the same form in every inertial reference system.

But this was not the only possible choice. It would also have been quite possible to have adopted the convention of referring all measurements to one particular inertial reference system such as that in which the solar system is at rest, i.e. of arbitrarily selecting this system as a "conventionally absolute" system replacing in practical measurements the elusive "absolute" system of the Maxwell-Lorentz theory.

There could be no logical objection to such a convention. Nor could there be any fundamental physical objection; the laws of electrodynamics had in fact been established in this system on the tacit, but probably mistaken, assumption that it was at rest in absolute space. Moreover, there would be no fundamental difficulty in referring all measurements to this conventionally absolute system. It had of course to be recognized that, in systems in conventionally absolute uniform motion, the measuring devices, such as rods and clocks, would be affected by the motion in accordance with the hypothesis of Fitzgerald and Lorentz and the entailed slowing down of clocks in motion (and in agreement with the predictions of the restricted theory).

There would in fact have been considerable conceptual advantages in such a convention. The relativity of simultaneity and the reciprocity of the relativistic variations of the restricted theory would not have obtruded, simply because the spatial and temporal coordinates of events in various inertial reference systems would not have been of conceptual interest. It would also have resulted in a description of the universe in terms of the measures of only one reference system and this would have avoided difficulties of definition such as that discussed in Section VII above.

Indeed, it can well be argued that such a convention would have been the obvious and sensible transition from the then-current ideas of absolute spatial and temporal measurements. For, as was shown elsewhere (Builder 1958), the concepts of absolute space and time remained essential in our physical description of the universe, while the fact that the absolute reference system could not be identified by observations of dynamical and electrodynamical phenomena (principle of relativity) could not in principle preclude its identification by purely geometrical and kinematical measurements; it has in fact always been believed, and with good reason, that the absolute velocity of the solar system, i.e. its velocity relative to the universe as a whole, must be very small indeed compared with the velocity of light.

The critical question to be answered is this: Would the adoption of such a convention have precluded the physical discoveries that are credited to the restricted theory?

It is clear that it could not have precluded such discoveries. The restricted theory itself, as defined in Section I, is a purely deductive inference from two postulates which are themselves statements, generalized from experience, of the physical characteristics of natural phenomena. The particular form of statement chosen for this inference, based on a particular convention about the synchronization of clocks, added to the postulates nothing more about the physical

characteristics of natural phenomena. Thus the restricted theory, by itself, could entail nothing about the characteristics of natural phenomena that was not already entailed by the postulates. Similarly, the restricted theory itself, taken together with the laws of electrodynamics and of dynamics, as modified by the postulates, could entail nothing about the characteristics of natural phenomena that was not already entailed by the postulates taken together with these laws.

It follows that the physical consequences deduced by means of the restricted theory could, from a purely logical viewpoint, have been equally well deduced using some alternative convention such as that suggested above. There is of course little doubt that the road to the discovery of these consequences would then have been longer and more arduous; no one would be inclined to dispute the elegance and effectiveness of the restricted theory formulation. We are, however, concerned here with assessing the fundamental character of the theory and not with assessing its elegance and its practical advantages.

We are, moreover, much concerned with the conceptual difficulties that have been caused by exponents of the restricted theory who have failed to recognize its conventional character and who have consequently asserted that its conventions demand a complete revision of our fundamental concepts of space and time.

Such exponents of the restricted theory will no doubt consider the point of view presented here as being, like the Fitzgerald-Lorentz hypothesis,* nothing but an "artful" device to rescue old concepts.

X. ACKNOWLEDGMENT

The author is indebted to Dr. C. A. Hurst of the Department of Mathematical Physics, University of Adelaide, for valuable critical comments in the course of this investigation.

XI. REFERENCES

- BUILDER, G. (1957).—The resolution of the clock paradox. *Aust. J. Phys.* **10**: 246.
 BUILDER, G. (1958).—Ether and relativity. *Aust. J. Phys.* **11**: 279.
 EDDINGTON, A. S. (1928).—"The Nature of the Physical World." (Cambridge Univ. Press.)
 EINSTEIN, A. (1905).—Elektrodynamik bewegter Körper. *Ann. Phys., Lpz.* (4) **17**: 891.
 EINSTEIN, A. (1907).—Über das Relativitätsprinzip. . . *Jb. Radioakt.* **4**: 411.
 GRÜNBAUM, A. (1955).—Logical and philosophical foundations of the special theory of relativity. *Amer. J. Phys.* **23**: 450.
 IVES, H. E. (1951).—Revisions of the Lorentz transformations. *Proc. Amer. Phil. Soc.* **95**: 125.
 LARMOR, J. (1900).—"Aether and Matter." (Cambridge Univ. Press.)
 LORENTZ, H. A. (1892).—The relative motion of the earth and the ether. *Versl. Gewone Vergad. Akad. Amst.* **1**: 74.
 LODGE, O. (1893).—Aberration problems. *Phil. Trans. A* **184**: 727.
 MATTHIAS, O. (1939).—Die ballistische Lichttheorie. . . *Phys. Z.* **40**: 443, 559.

* "Diese ad hoc eingeführte Annahme erschien aber doch nur als ein künstliches Mittel, um die Theorie zu retten" (Einstein 1907, pp. 412–13).

- MINKOWSKI, H. (1908).—"Space and Time." Translation by Perrett and Jeffery: "The Principle of Relativity." (1923.) (Methuen: London.)
- POINCARÉ, H. (1904).—L'état actuel et l'avenir de la physique mathématique. *Bull. Sci. Math.* **28**: 302. (English translation. *Monist* **15**: 1 (1905).)
- RITZ, W. (1908).—Recherches critiques sur l'électrodynamique générale. *Ann. Chim. (Phys.)* (8) **13**: 145.
- SILBERSTEIN, L. (1922).—"The Theory of General Relativity and Gravitation." (Van Nostrand: New York.)
- SILBERSTEIN, L. (1924).—"The Theory of Relativity." (Macmillan: London.)
- SOMMERFELD, A. (1952).—"Electrodynamics." (Academic Press: New York.)
- WHITTAKER, E. (1953).—"A History of the Theories of Aether and Electricity: 1900-1926." (Thomas Nelson: London.)

ELECTRON EXCITATION OF COLLECTIVE NUCLEAR TRANSITIONS

By L. J. TASSIE*

[Manuscript received July 10, 1958]

Summary

The inelastic scattering of high energy electrons with excitation of collective nuclear transitions is treated using a simple hydrodynamical model in which the collective nuclear motion is assumed to be irrotational and incompressible. The effects of nuclear compressibility are discussed. Using the Born approximation, scattering form factors are calculated for several charge distributions for electric quadrupole transitions, and the sensitivity of the scattering to the form of the nuclear charge distribution is examined.

I. INTRODUCTION

In a previous paper (Tassie 1956) the inelastic scattering of high energy electrons by nuclei was considered using a modified liquid drop model of the nucleus which allows for non-uniform nuclear charge and mass density distributions. The collective nuclear motion was assumed to be irrotational and incompressible, and the Born approximation was used to calculate the electron scattering using several simple forms for the nuclear charge distribution.

It is the purpose of the present paper to discuss the incompressibility assumption, to extend the calculations of the inelastic electron scattering using more realistic forms for the nuclear charge distribution, and to examine how sensitive this scattering is to the choice of the nuclear charge distribution.

II. GENERAL THEORY

The differential cross section for the inelastic scattering of high energy electrons with excitation of a nuclear electric 2^L -pole transition is (Schiff 1954)

$$d\sigma/d\omega = (d\sigma/d\omega)_p |\mathcal{F}_L|^2, \dots\dots\dots (1)$$

where $(d\sigma/d\omega)_p = \frac{1}{4}Z^2(e^2/\hbar c)^2 k^{-2} \cos \frac{1}{2}\theta \operatorname{cosec}^4 \frac{1}{2}\theta$ is the point charge scattering cross section, and

$$\mathcal{F}_L = 4\pi(2L+1)^{\frac{1}{2}} \int j_L(Qr) Y_{L,0} \rho_{\text{trans}} dV \dots\dots\dots (2)$$

is the nuclear form factor. $Z\rho_{\text{trans}}$ is the transition charge density of the nucleus, $\hbar k$ is the momentum of the incident electron, and $q = 2k \sin \frac{1}{2}\theta$.

For the nuclear model used here, equation (2) becomes (Tassie 1956)

$$\mathcal{F}_L = 2\pi^{\frac{1}{2}} Q_{L,0} I_L / Z e (2L+1)^{\frac{1}{2}} r^{2(L-1)} \dots\dots\dots (3)$$

where

$$Q_{L,0} = Z e \int r^L Y_{L,0} \rho_{\text{trans}} dV \dots\dots\dots (4)$$

* Research School of Physical Sciences, Australian National University, Canberra.

is the nuclear transition 2^L -pole moment, and

$$I_L = 4\pi q \int_0^\infty j_{L-1}(qr) r^{L+1} \rho(r) dr. \quad \dots\dots\dots (5)$$

$Ze\rho(r)$ is the static charge density of the nucleus. The angular dependence of \mathcal{F}_L is then given completely by I_L .

For elastic scattering by a spherically symmetrical nucleus, the cross section is given by equation (1) with

$$\mathcal{F}_{\text{elastic}} = 4\pi \int_0^\infty j_0(qr) \rho(r) r^2 dr. \quad \dots\dots\dots (6)$$

Then equation (5) can be written

$$I_L = (-1)^{L+1} q^{L-1} \left(\frac{\partial}{\partial q} q^{-1} \right)^{L-2} \frac{\partial}{\partial q} \mathcal{F}_{\text{elastic}}, \quad \dots\dots\dots (7)$$

and in particular, for electric quadrupole ($E2$) transitions we have

$$I_2 = -q \partial \mathcal{F}_{\text{elastic}} / \partial q. \quad \dots\dots\dots (8)$$

As $q \rightarrow 0$,

$$I_2 \rightarrow \frac{1}{2} q^2 \langle r^2 \rangle, \quad \dots\dots\dots (9)$$

where

$$\langle r^2 \rangle = 4\pi \int_0^\infty \rho(r) r^4 dr. \quad \dots\dots\dots (10)$$

The values of $\langle r^2 \rangle$ are required so that the transition electric quadrupole moments can be determined by comparing I_2 with experimental results for the electron scattering form factors.

The above results are obtained by using the Born approximation, and the reliability of this must be examined. The plane wave of the incident electron is distorted by the Coulomb field of the nucleus arriving at the nucleus looking very much like a plane wave with modified amplitude and wave number and slightly curved wave fronts. Downs, Ravenhall, and Yennie (1957) have considered the inelastic scattering using a perturbation method which includes the effect of this distortion, and their results are similar to the Born approximation results obtained using a slightly modified q and with a partial filling in of the Born approximation diffraction minima. A change in q in equations (5) or (6) is equivalent to a change in the size of the nucleus, so that the effect of the modification of q can be taken into account by correcting nuclear size parameters which have been derived from experiment using the Born approximation. This procedure has been used by Fregeau (1956) and by Helm (1956).

III. NUCLEAR COMPRESSIBILITY

The effects of nuclear compressibility in the liquid drop model have been investigated assuming the nucleus has a sharp surface (Woeste 1952). However, since the nucleus has a diffuse surface (Hofstadter 1956), the energy density of the nuclear fluid must depend on the derivatives of the nuclear density (Swiatecki

1950), and this leads to complicated hydrodynamical equations. The problem is considerably simplified by neglecting compressibility, so that

$$\nabla^2\Phi=0, \quad \dots\dots\dots (11)$$

where Φ is the velocity potential of the nuclear fluid, and the results of Section I are then obtained.

Another approach is to rely on the sharp-edged liquid drop model for an estimate of the effect of compressibility, and to use this estimate for a nucleus with a diffuse edge. Then compressibility can be neglected if the velocity of the fluid is everywhere small compared to the velocity of "sound" in the nucleus. For quadrupole vibrations satisfying (11), the maximum velocity is given by

$$(v_{\max.}/C_s)^2=(5/3)\hbar\omega/AmC_s^2, \quad \dots\dots\dots (12)$$

and occurs at the edge of the nucleus. $\hbar\omega$ is the one-phonon energy and C_s is the velocity of sound in the nuclear fluid. Using the hydrodynamical estimates of $\hbar\omega$ given by Bohr and Mottelson (1953) and the value given by Rosenfeld (1948) for C_s , we obtain

$$(v_{\max.}/C_s)^2\simeq 6\cdot 5A^{-3/2}. \quad \dots\dots\dots (13)$$

Thus, for very light nuclei the incompressible approximation is not justified, but a collective treatment is not strictly applicable to light nuclei. For heavy nuclei, equation (11) gives some justification to the use of the incompressible approximation, e.g. for $A=16$, $(v_{\max.}/C_s)^2\simeq 0\cdot 1$.

The possible indirect effect of compressibility on the scattering form factor by its effect on the static charge distribution of the nucleus is discussed at the end of Section IV.

IV. CALCULATIONS

The following functional forms have been used for the nuclear charge distribution :

Fermi :

$$\rho_F(r)=\rho_{0F}/\{\exp [(r-c_F)/z_F]+1\}, \quad \dots\dots\dots (14)$$

Trapezoidal :

$$\rho_T(r)=\rho_{0T}, \quad \left. \begin{array}{l} r \leq c_T - z_T, \\ = \rho_{0T}(c_T + z_T - r)/2z_T, \quad c_T - z_T \leq r \leq c_T + z_T, \\ = 0, \quad r \geq c_T + z_T. \end{array} \right\} \dots (15)$$

Modified Gaussian :

$$\rho_G(r)=\rho_{0G}/\{\exp [(r^2-c_G^2)/z_G^2]+1\}. \quad \dots\dots\dots (16)$$

Three-parameter :

$$\rho_M(r)=\rho_{0M}[1+(wr^2/c_M^2)]/\{\exp [(r-c_M)/z_M]+1\}. \quad \dots (17)$$

Hahn, Ravenhall, and Hofstadter (1956) have used these forms for the charge distribution in analysing elastic scattering.

Blankenbecler (1957) has obtained an approximate expression for the Born approximation elastic scattering form factor for a Fermi charge distribution. The correction to this can be obtained as a power series in $\exp(-c_F/z_F)$, so that we finally obtain

$$\begin{aligned} \mathcal{F}_{c_F} = & 4\pi^2 \rho_{0F} c_F z_F q^{-1} \operatorname{cosech} \pi q z_F \\ & \times (\pi z_F c_F^{-1} \sin q c_F \cotanh \pi q z_F - \cos q c_F) \\ & + 8\pi \rho_{0F} z_F^3 \sum_{n=1}^{\infty} (-1)^{n-1} n \exp[-n c_F/z_F] / (n^2 + q^2 z_F^2)^2, \dots \dots \dots (18) \end{aligned}$$

$$\rho_{0F} = (3/4\pi c_F^3) \{1 + (\pi z_F/c_F)^2 - \frac{1}{3}(z_F/c_F)^3 \sum_{n=1}^{\infty} (-1)^{n-1} \exp[-n c_F/z_F] / n^3\}^{-1}. \dots (19)$$

Using equation (8), we obtain

$$\begin{aligned} I_{2F} = & 4\pi^2 \rho_{0F} q^2 c_F^5 \\ & \times \{(z_F/c_F) \operatorname{cosech} \pi q z_F [(q c_F)^{-3} (\pi z_F c_F^{-1} \sin q c_F \cotanh \pi q z_F - \cos q c_F) \\ & - (q c_F)^{-2} (\sin q c_F (1 - \pi^2 z_F^2 c_F^{-2} - 2\pi^2 z_F^2 c_F^{-2} \operatorname{cosech}^2 \pi q z_F) \\ & + 2\pi z_F c_F^{-1} \cotanh \pi q z_F \cos q c_F)] \\ & + 8\pi^{-1} z_F^5 c_F^{-5} \sum_{n=1}^{\infty} (-1)^{n-1} n \exp[-n c_F/z_F] (n^2 + q^2 z_F^2)^{-3}\}. \dots \dots (20) \end{aligned}$$

Also*

$$\begin{aligned} \langle r^2 \rangle = & 4\pi \rho_{0F} c_F^5 \{ [3 + 10(\pi z_F/c_F)^2 + 7(\pi z_F/c_F)^4] / 15 \\ & + 120(z_F/c_F)^2 \sum_{n=1}^{\infty} (-1)^{n-1} n^{-5} \exp[-n c_F/z_F] \} \dots (21) \end{aligned}$$

$$\simeq 3c_F^2 \{1 + 7(\pi z_F/c_F)^2 / 3\} / 5. \dots \dots \dots (22)$$

The corresponding expressions for the trapezoidal charge distribution are easily obtained:

$$\rho_{0T} = 3\{4\pi c_T^3 [1 + (z_T/c_T)^2]\}^{-1}, \dots \dots \dots (23)$$

$$\begin{aligned} I_{2T}(q) = & 4\pi \rho_{0T} (z_T q^4)^{-1} \\ & \times \{ -(q z_T)^2 \sin q z_T \sin q c_T \\ & + \sin q z_T [(8 - (q c_T)^2) \sin q c_T - 5 q c_T \cos q c_T] \\ & + q z_T \cos q z_T [2 q c_T \cos q c_T - 5 \sin q c_T] \}, \dots (24) \end{aligned}$$

$$\langle r^2 \rangle = (4\pi \rho_{0T} c_T^5 / 15) [1 + 3(z_T/c_T)^2] [3 + (z_T/c_T)^2]. \dots (25)$$

The three distributions, equations (12), (13), (14), reduce to the uniform charge distribution for $z=0$ and then

$$I_2(q) = 3j_2(qc). \dots \dots \dots (26)$$

The mean square radii, $\langle r^2 \rangle$, are given in Table 1, and Figures 1 and 2 show the values of I_2 for the Fermi and trapezoidal distributions respectively. The

* A derivation of equation (21) is given by Elton (1958), but the expression itself is stated incorrectly. The correction of this mistake slightly alters the nuclear radius parameters obtained by Elton but does not affect his conclusions (Elton, personal communication).

ratio of the inelastic scattering to point charge scattering, which is proportional to $|I_2|^2$, will consist of a series of diffraction maxima separated by zeros. The main effect of increasing the thickness of the nuclear surface is to decrease the higher order diffraction maxima.

The zeros of $|I_2|^2$ are due to the use of the first Born approximation. In a more accurate treatment, in which I_2 is complex, these zeros would become diffraction minima (Downs, Ravenhall, and Yennie 1957). In general, the excited state of the nucleus will not be completely described by collective excitation, but will also consist of some single-particle excitation as in the Bohr-Mottelson collective model (Bohr and Mottelson 1953). This admixture

TABLE 1

$$\langle r^2 \rangle = 4\pi \int_0^\infty \rho(r) r^4 dr$$

	Fermi Charge Distribution (Equation (14))					
z_F/c_F	0	0.08	0.16	0.2060	0.1287	0.0839
$\langle r^2 \rangle / c_F^2$	0.6000	0.6884	0.9537	1.1856	0.8289	0.6971
	Trapezoidal Charge Distribution (Eqn. (15))			Modified Gaussian Distribution (Eqn. (16))	Three-Parameter Distribution (Eqn. (17))	
z/c	0.2373	0.4745	0.7504	0.5657	0.1010	
w	—	—	—	—	0.64	
$\langle r^2 \rangle / c^2$	0.6764	0.8822	1.2261	0.8501	0.8204	

of single-particle excitation can also be expected to fill in the diffraction minima to some extent. However, the corrections to the Born approximation and the admixture of single-particle excitation should not greatly affect the magnitude of the diffraction maxima, and, for this reason, a useful quantity for comparison with experiment should be the ratio (*second maximum of $|I_2|^2$*)/(*first maximum of $|I_2|^2$*) = $(M_2/M_1)^2$. M_2 is the first non-zero minimum of I_2 , and M_1 is the first maximum of I_2 .

M_2/M_1 has been obtained by using the above results for I_2 for the Fermi and trapezoidal charge distributions, and is shown in Figure 3 as a function of $(z/c)^2$. Hahn, Ravenhall, and Hofstadter (1956) have used a parameter t , the surface thickness, which is defined as the distance between the points where ρ has 0.9 and 0.1 of its maximum value. The scales of $(z_F/c_F)^2$ and $(z_T/c_T)^2$ in Figure 3 have been chosen so that the figure also shows M_2/M_1 as a function of $(t/c)^2$. For this purpose the relations, $t = 1.60z_T$, $t = (4 \ln 3)z_F = 4.40z_F$,* have been used.

* To be exact

$$t = \{4 \ln 3 + \ln [1 - (80/9)(10 + \exp(c_F/z_F))^{-1}]\} z_F.$$

The scale of $(t/c)^2$ for the Fermi distribution has a maximum error of 2 per cent.

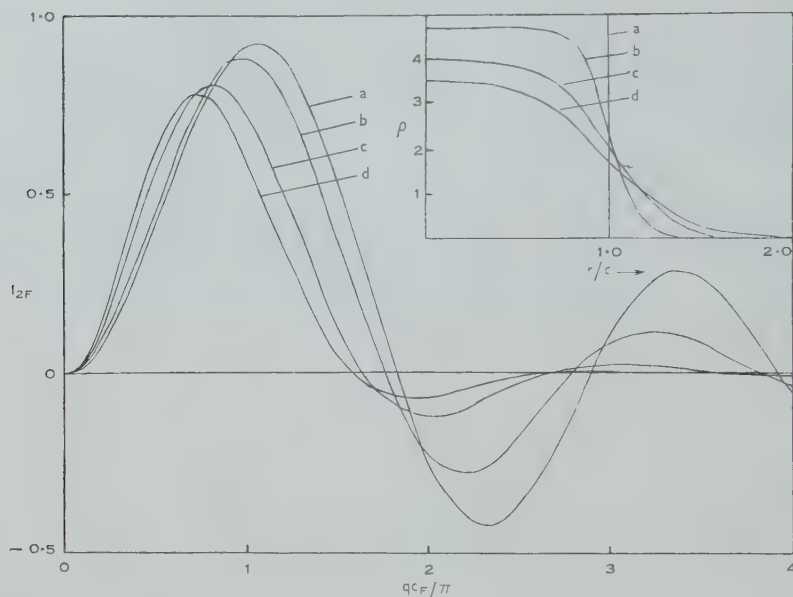


Fig. 1.— $I_{2F}(q)$, equation (5), for the Fermi charge distribution, equation (14), for the following values of z_F/c_F : (a) 0, (b) 0.08, (c) 0.16, (d) 0.206. The charge density distributions for these values of z_F/c_F are also shown. For $z_F/c_F=0$, the Fermi distribution becomes the uniform charge distribution and then $I_{2F}(q)=3j_2(qc_F)$, curve a.

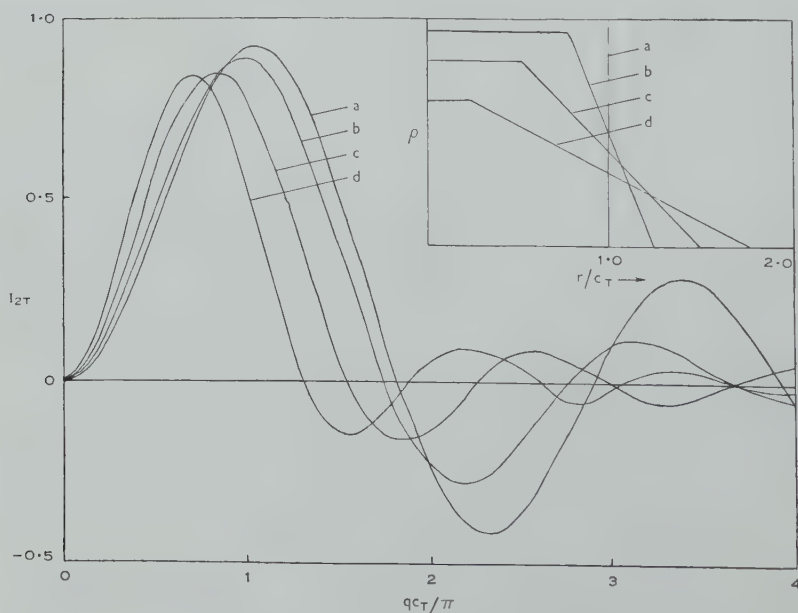


Fig. 2.— $I_{2T}(q)$, equation (5), for the trapezoidal charge distribution, equation (15), for the following values of z_T/c_T : (a) 0, (b) 0.2373, (c) 0.4745, (d) 0.7504. The charge density distributions for these values of z_T/c_T are also shown. For $z_T/c_T=0$, the trapezoidal distribution becomes the uniform charge distribution and then $I_{2T}(q)=3j_2(qc_T)$, curve a.

The figure indicates that I_2 is to some extent sensitive to the functional form of ρ , more especially for large values of t . It is seen that it is impossible to obtain small values of M_2/M_1 for the trapezoidal charge distributions.

I_2 for the modified Gaussian distribution, equation (16), and the three-parameter distribution, equation (17), have been obtained by numerical integration of equation (5). Figure 4 shows I_{2G} for $(z_G/c_G)^2=0.32$ compared with the $I_{2F}(z_F/c_F=0.1287)$ which has the same value of M_2/M_1 . The scale of qc_F/π has been chosen to match the second zeros of the two curves. Thus, in

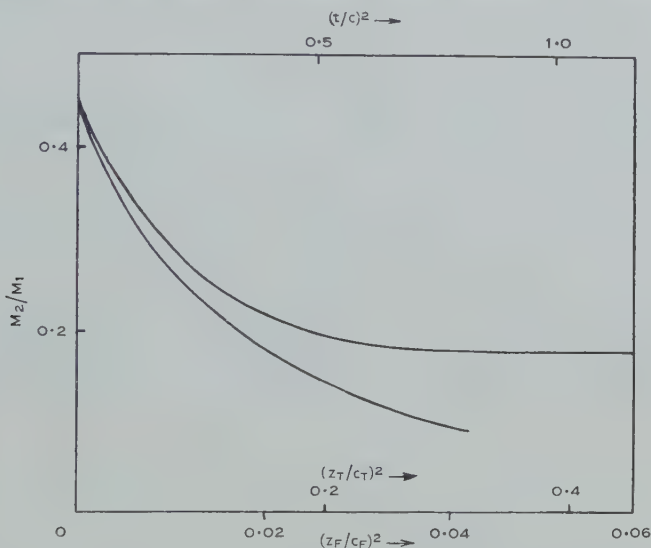


Fig. 3.— M_2/M_1 for the Fermi charge distribution, equation (14), as a function of $(z_F/c_F)^2$, and for the trapezoidal charge distribution, equation (15), as a function of $(z_T/c_T)^2$. The scales of $(z_F/c_F)^2$ and $(z_T/c_T)^2$ are chosen so that the figure gives M_2/M_1 as a function of $(t/c)^2$. t is the surface thickness, which is defined as the distance between the points where ρ has 0.9 and 0.1 of its maximum value. M_2 is the first non-zero minimum of I_2 . M_1 is the first maximum of I_2 .

Figure 4 the scattering form factor of a nucleus with a modified Gaussian charge distribution and a transition quadrupole moment Q_{L0} is compared with the scattering form factor for a nucleus with a Fermi charge distribution with radius $c_F=c_G/0.95$, and a quadrupole moment $Q_{L0}/(0.95)^3$.

I_{2M} for the three-parameter charge distribution for $z_M/c_M=0.1010$ and $w=0.64$ is compared in Figure 5 with I_{2F} for the Fermi distribution with $z_F/c_F=0.0839$, taking $c_F=1.052c_M$. Hahn, Ravenhall, and Hofstadter (1956) have made accurate calculations of the elastic scattering by these two charge distributions with these values of the parameters and show that the experimental elastic scattering cannot distinguish between these two charge distributions. From Figure 5 it is seen that it would be very difficult to distinguish between these two charge distributions by measurement of the collective electric quadrupole inelastic scattering.

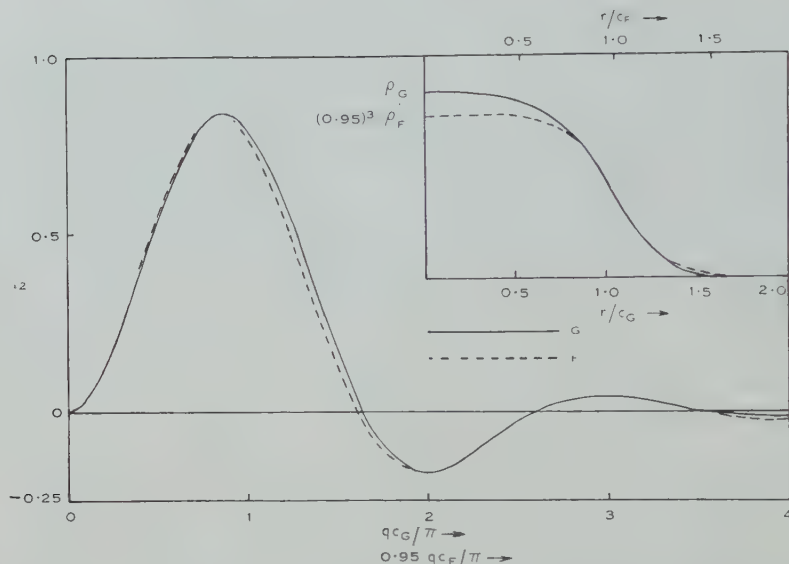


Fig. 4.— $I_{2G}(q)$, equation (5), for the modified Gaussian charge distribution, equation (16), $(z_G/c_G)^2=0.32$, curve G , compared with $I_{2F}(q)$ for the Fermi charge distribution, equation (14), with $z_F/c_F=0.1287$. The two charge distributions are also shown.

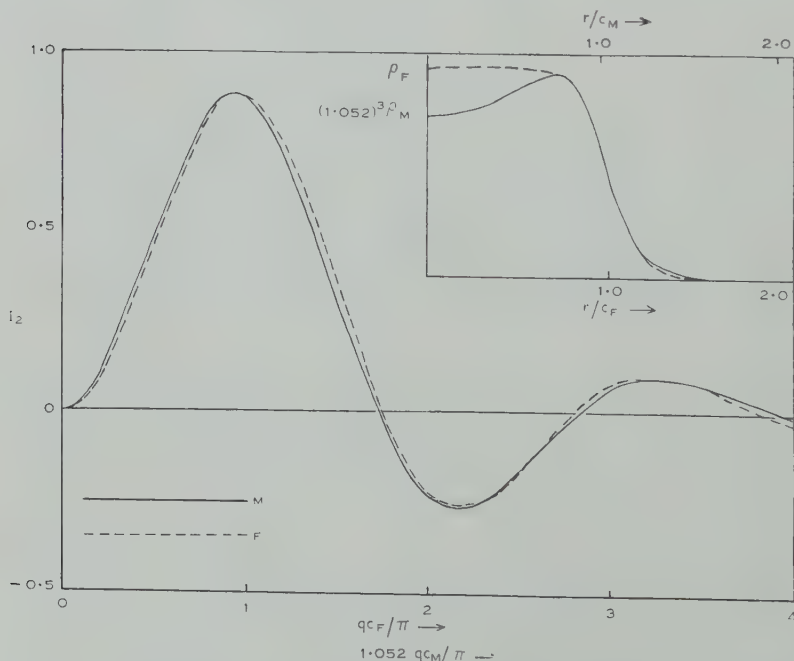


Fig. 5.— $I_{2M}(q)$, equation (5), for the three-parameter charge distribution, equation (17), with $z_M/c_M=0.1010$ and $w=0.64$, curve M , compared with $I_{2F}(q)$ for the Fermi charge distribution, equation (14), with $z_F/c_F=0.0839$. The two charge distributions are also shown.

The situation described by the three-parameter charge distribution, with the charge density tending to be smaller in the middle of the nucleus, is expected to arise from the effects of nuclear compressibility (Woeste 1952). Since this static effect of nuclear compressibility causes only a small modification to $I_2(q)$, it seems safe to neglect this effect of compressibility at least for the range of qc considered here.

The mean square radii of the various charge distributions are given in Table 1.

V. DISCUSSION

The inelastic scattering form factor is not sensitive to the form of the charge density distribution except for nuclei with very diffuse surfaces, i.e. for large values of the surface thickness. The most satisfactory test of the theory given here would be the experimental determination of (M_2/M_1) , the ratio of the second and first peaks in $|\mathcal{F}_2|$. Helm (1956) has performed experiments on the electron excitation of several electric quadrupole ($0^+ \rightarrow 2^+$) nuclear transitions, but his results are not sufficiently comprehensive to allow the determination of $(M_2/M_1)^2$. Other effects, such as the possibility of some single-particle excitation, must be considered in more detail before comparing the theoretical and experimental angular distributions of the inelastically scattered electrons.

VI. ACKNOWLEDGMENTS

Numerical integrations were performed on CSIRAC (University of Melbourne), and the author wishes to thank the computer staff for their assistance. The remaining computations were carried out by Mrs. B. Nerdal, to whom the author expresses his thanks. The author is also indebted to Professor K. J. Le Couteur for his interest in this work.

VII. REFERENCES

- BLANKENBECLER, R. (1957).—*Amer. J. Phys.* **25**: 279.
 BOHR, A., and MOTTELSON, B. R. (1953).—*Math. fys. Medd.* **27**: No. 16.
 DOWNS, B. W., RAVENHALL, D. G., and YENNIE, D. R. (1957).—*Phys. Rev.* **106**: 1285.
 ELTON, L. R. B. (1958).—*Nuclear Phys.* **5**: 173.
 FREGEAU, J. H. (1956).—*Phys. Rev.* **104**: 225.
 HAHN, B., RAVENHALL, D. G., and HOFSTADTER, R. (1956).—*Phys. Rev.* **101**: 1131.
 HELM, R. H. (1956).—*Phys. Rev.* **104**: 1466.
 HOFSTADTER, R. (1956).—*Rev. Mod. Phys.* **28**: 214.
 ROSENFELD, L. (1948).—"Nuclear Forces." (North Holland Publishing Co.: Amsterdam.)
 SCHIFF, L. I. (1954).—*Phys. Rev.* **96**: 765.
 SWIATECKI, W. J. (1950).—*Proc. Phys. Soc. A* **63**: 1208.
 TASSIE, L. J. (1956).—*Aust. J. Phys.* **9**: 407.
 WOESTE, K. (1952).—*Z. Phys.* **133**: 370.

THE PHOTODISINTEGRATION OF NUCLEI WITH Z BETWEEN 9 AND 30

By B. M. SPICER*

[Manuscript received July 14, 1958]

Summary

It is pointed out that the Danos-Okamoto effect (splitting of the giant resonance for non-spherical nuclei into two components) should be more readily observable in the region $9 \leq Z \leq 30$ than in the rare earth region where the initial search for the effect has been made. The ratio of the energies of the two components of the split resonance depends on the deformation of the nucleus from spherical shape. Deformations are even larger in this region of atomic number than for the rare earth nuclei, whose intrinsic electric quadrupole moments are large. The predictions of the theory of Danos (1958) are shown to be verified experimentally in five cases. Five other cases where the effect should be observable are discussed.

I. INTRODUCTION

The application to non-spherical nuclei of the long-range correlation model of the nuclear photoeffect was made by Danos (1958) and Okamoto (1956, 1958). Their work predicted a splitting of the giant resonance into two components for these nuclei. If $\hbar\omega_a$ and $\hbar\omega_b$ are the energies of the two resonances, they are related to the shape of the nucleus in its ground state by

$$\hbar\omega_b/\hbar\omega_a = 0.911a/b + 0.089, \quad \dots\dots\dots (1)$$

where a and b are the lengths of the half axes of the nucleus, which is assumed spheroidal. a refers to the axis of rotational symmetry. The electric quadrupole moment of a spheroid having a uniform charge distribution is given by

$$Q_0 = \frac{2}{5}Z(a^2 - b^2). \quad \dots\dots\dots (2)$$

Defining $R^3 = R_0^3 A = ab^2$, where R is the radius of a sphere having the same volume as the spheroid, we can write the eccentricity ε as $(a^2 - b^2)/R^2$. This gives

$$Q_0 = \frac{2}{5}R_0^2\varepsilon \cdot ZA^{2/3}. \quad \dots\dots\dots (3)$$

The electron scattering of Hofstadter (1956), when interpreted in terms of a uniform charge distribution, leads to $R_0 = 1.2$ fermi. That value is used throughout this paper.

If the eigenvalue splitting ($\hbar\omega_b - \hbar\omega_a$) is less than Γ (the width at half maximum cross section of the giant resonance for spherical nuclei), the cross section consists of one broadened peak, but, if the splitting is greater than Γ , then two peaks should be resolved if their heights are approximately equal.

* Physics Department, University of Melbourne.

In a nucleus with a positive Q_0 , the peak of greater integrated cross section occurs at higher energy, while the opposite is true if the nucleus concerned has a negative Q_0 .

This theory has been verified experimentally for tantalum and terbium by Weiss and Fuller (1958), and independently by Spicer, Thies, *et al.* (1958) for tantalum. Both these experiments indicate the existence of two peaks in the cross section for nuclear absorption of photons, although neither experiment indicated two completely separated peaks.

Since the theory of the Danos-Okamoto effect is based on a strong interaction model of the nucleus, all nuclei where the collective model (Bohr and Mottelson 1953) has been applied successfully should also exhibit the splitting of the giant resonance. The purpose of the present paper is to point out that the range of nuclei $9 \leq Z \leq 30$ is one in which the Danos-Okamoto effect should be observable.

In this region of atomic number, the Chalk River group (Litherland *et al.* 1956; Bromley, Gove, and Litherland 1957; Litherland *et al.* 1958) have been able to account for a large body of experimental data in terms of one unique spheroidicity parameter δ for each nucleus considered. This spheroidicity parameter δ was defined by Nilsson (1955) and is related to the ϵ of Danos by

$$\epsilon = 2\delta(1 + \frac{2}{3}\delta).$$

This expression was obtained from comparison of the expressions for Q_0 given by Nilsson (1955) and Danos (1958).

Therefore, a number of nuclei in the range $9 \leq Z \leq 30$ are examined for the existence of the Danos-Okamoto effect. Using values of Q_0 obtained from measurements in either Coulomb excitation or microwave spectroscopy, the ratio of the two resonance energies for each nucleus is calculated, and, where possible, compared with experiment. Table 1 summarizes this comparison, and discussion follows below.

TABLE 1
COMPARISON OF PREDICTED AND OBSERVED RESONANCE ENERGY RATIOS

Nucleus	ϵ	Reference	$\hbar\omega_b/\hbar\omega_a$	
			Predicted	Observed
^{19}F	0.69	Sherr, Li, and Christy (1955)	1.31	1.43
^{26}Mg	0.89	Litherland <i>et al.</i> (1958)	1.43	1.45
$^{29,30}\text{Si}$	-0.37	Bromley, Gove, and Litherland (1957)	0.84	0.73
^{51}V	0.28	Alder <i>et al.</i> (1956)	1.12	1.17
^{63}Cu	-0.30	Bleaney, Bowers, and Pryce (1955)	0.86	0.87

The resonances found on the low energy side of the giant resonance for two of the nuclei considered, namely ^{19}F and ^{26}Mg , were formerly classed as "pigmy" resonances. They are here reinterpreted as components of a split giant resonance.

II. EXAMINATION OF DATA

The ratio of the two resonance energies as given by Danos (1958) is linearly dependent on the ratio a/b . Verification of the existence of this effect was first sought in or near the rare earth region, since it is there that large electric quadrupole moments are found. Since the difference between the two resonance energies is a function of a/b and since the factor $ZA^{\frac{1}{3}}$ appears in the expression for Q_0 , it was thought worthwhile to seek through other regions of the periodic table for nuclei having large values of ϵ , and hence of a/b . The region $9 \leq Z \leq 30$ is a region where ϵ -values are found of the same magnitude or larger than those in the rare earth region.

(a) ^{19}F ((γ, n) threshold $10.41 \pm 0.01 \text{ MeV}$; (γ, p) threshold $7.96 \pm 0.01 \text{ MeV}$)

It has been pointed out by Paul (1957) and Rakavy (1957) that the energy level structure of the low-lying levels of ^{19}F can be accounted for satisfactorily in terms of the collective model. This leads to a deformation parameter δ of 0.3. This interpretation also gives correctly the ratio of γ -ray transition probabilities from the $3/2^+$ -state (1.59 MeV) to the ground and second excited states, the magnetic moments for the ground and second excited states, and the $E2$ lifetime for the second excited state. The above value of δ gives an ϵ of 0.72.

Coulomb excitation of the 198-keV level of ^{19}F by α -particles leads to a value of $B(E2)$ for the electric quadrupole excitation of this level. $B(E2)$ is the reduced transition probability for the γ -ray transition, and is defined by Bohr and Mottelson (1953). It is calculable from the yield function for γ -rays from the Coulomb excitation process, and in this case is found to be $0.0034 \times 10^{-48} \text{ e}^2 \text{ cm}^4$ (Sherr, Li, and Christy 1955).

$B(E2)$ is related to the intrinsic quadrupole moment for the nucleus (Heydenburg and Temmer 1956), in this case by the relation

$$Q_0^2 = \frac{8}{15} \pi (2I_0 + 3)(I_0 + 2) B(E2),$$

which holds for $E2$ transitions I_0 to $I_0 + 1$, I_0 being the spin of the initial state. The quoted value of $B(E2)$ leads to a value of 0.24 barn for Q_0 . From this value of Q_0 , whose absolute accuracy is estimated as about 25 per cent., we find ϵ to be 0.65, in good agreement with the value obtained by Paul (1957). The mean value of ϵ gives a/b as 1.34. The ratio of the two resonance energies $\hbar\omega_b/\hbar\omega_a$ for the photodisintegration process is thus predicted to be 1.31. If we assume $\hbar\omega_b$ to be 20 MeV, then $\hbar\omega_a$ is predicted to be 15.1 MeV.

A measurement of the $^{19}\text{F}(\gamma, n)^{18}\text{F}$ cross section by Horsley, Haslam, and Johns (1952) shows evidence for a peak on the low energy side of the 20 MeV resonance. The position of this lower energy peak is uncertain but is estimated as 14 MeV. We interpret these peaks as the two components of the split giant resonance due to the non-spherical shape of ^{19}F . Both peaks must then be due to electric dipole transitions.

The (γ, p) cross section is assumed to have approximately the same shape as the (γ, n) , allowing for the (γ, p) threshold being 2.4 MeV lower than the

(γ, n) threshold. Lasich, Muirhead, and Shute (1955) measured angular distributions of photoprotons from fluorine for γ -ray energies between 10 and 17 MeV. They concluded that the multipolarity of the radiation absorbed in this energy region is predominantly either electric dipole or magnetic dipole. The $E1$ possibility is required by the interpretation of this paper.

(b) ^{25}Mg ((γ, n) threshold, 7.33 ± 0.02 MeV; (γ, p) threshold, 12.07 ± 0.02 MeV)

The properties of ^{25}Mg are discussed in terms of the collective model by Gove *et al.* (1956) and by Litherland *et al.* (1958). These latter authors conclude, on the basis of a detailed analysis of the properties of ^{25}Mg and ^{25}Al , that the equilibrium distortion of both these nuclei lies between $\eta=4$ and $\eta=6$, with the value 4.5 most favoured. η (see Nilsson 1955) is related to δ by

$$\eta = \frac{\delta}{\kappa} \left[1 - \frac{4}{3}\delta^2 - \frac{16}{27}\delta^3 \right]^{-\frac{1}{2}}.$$

κ is a measure of the strength of the spin-orbit interaction assumed by Nilsson in the calculation of particle states in a spheroidal potential. He obtained a value of $\kappa=0.05$ for nuclei of atomic mass greater than 100. For the atomic mass region considered, the value suggested is 0.08 (see Bromley, Gove, and Litherland 1957). This lack of accurate knowledge of κ thus limits our knowledge of δ .

Using $\kappa=0.08$ and taking η as 4.5, we get $\delta=0.36$. The value of ε corresponding to this value of δ is 0.89. Following the same procedure as for ^{19}F , the ratio of the two resonance energies for photon absorption is 1.43. This assumes a positive electric quadrupole moment and is reasonable since surrounding nuclei (^{23}Na and ^{27}Al) are known to have positive quadrupole moments. If we take $\hbar\omega_b$ as 20 MeV, then $\hbar\omega_a$ is given as 14 MeV.

Nathans and Yergin (1955) determined the cross section for the $^{25}\text{Mg}(\gamma, n)$ reaction using a target of magnesium oxide enriched to 92.33 per cent. in ^{25}Mg . They found that this cross section exhibited a plateau in the region $13\frac{1}{2}$ –15 MeV and showed the expected giant resonance at 20 MeV. However, Katz *et al.* (1954) have found a peak at 13.5 MeV, and Spicer, Allum, *et al.* (1958) at 13.8 MeV in the cross section computed from the yield curve of neutrons from natural magnesium. This peak is believed to be due mainly to neutrons emitted from ^{25}Mg . The (γ, n) thresholds in ^{24}Mg and ^{26}Mg are 16.57 ± 0.02 and 11.12 ± 0.03 MeV respectively. ^{24}Mg cannot make any contribution to the 13.5 MeV peak on energetic grounds. ^{26}Mg has its (γ, n) threshold some 4 MeV above that of ^{25}Mg and would therefore be expected to contribute much less because of this. The discussion of the present paper requires that the 13.8 MeV peak be due to electric dipole absorption of photons. The 20 MeV peak is also expected to be due to $E1$ transitions, since it comes in the region expected for the giant dipole resonance.

Spicer, Allum, *et al.* (1958) find the width of the 13.8 MeV resonance to be 2 MeV. Since Nathans and Yergin (1955) give the only unambiguous data on the 20 MeV resonance in the (γ, n) cross section, their observed width is noted as

5 MeV. Toms and Stephens (1951), measuring the $^{25}\text{Mg}(\gamma, p)$ cross section, find a resonance at 20.5 MeV, with width of 3.5 MeV. The energy of the (γ, p) threshold is high enough to prevent the 13.8 MeV peak appearing in the cross section of this reaction. Thus these widths are in approximate agreement with those found necessary to fit the measured cross section for photon absorption in tantalum (Spicer, Thies, *et al.* 1958).

If the 13.8 MeV peak found by Spicer, Allum, *et al.* (1958) is ascribed to ^{25}Mg alone, then it is approximately the same height as the 20 MeV peak. This latter height was obtained by adding the peak cross sections found by Nathans and Yergin (γ, n) and Toms and Stephens (γ, p) . Thus the condition required by Danos's theory that $\int \sigma dE$ for the higher energy peak shall be twice that for the lower energy peak is approximately satisfied.

The interpretation given here is consistent with the available experimental data, although the reason for the non-appearance of the 13.8 MeV peak in the results of Nathans and Yergin is still unknown.

(c) ^{29}Si ((γ, n) threshold, 8.48 ± 0.03 MeV; (γ, p) threshold, 12.34 ± 0.03 MeV)

^{30}Si ((γ, n) threshold, 10.61 ± 0.03 MeV; (γ, p) threshold, 13.80 ± 0.11 MeV)

^{29}Si has been the subject of a detailed examination to see whether its properties can be correlated with the predictions of the collective model (Bromley, Gove, and Litherland 1957). These authors conclude that the correlation can be made, and assign the parameter δ a value of -0.22 (assuming $\kappa=0.08$ as for ^{25}Mg). This value of δ corresponds to ε equals -0.37_4 . This leads to a value for the ratio $\hbar\omega_b/\hbar\omega_a$ of 0.836. Since the minus sign of δ implies an oblate spheroid, the peak of larger integrated cross section must come at the lower energy. Taking $\hbar\omega_a$ as 20 MeV, $\hbar\omega_b$ is predicted as 16.7 MeV.

Evidence from photodisintegration reactions is given by Katz *et al.* (1954). We assume that what is obtained by Katz for the low abundance ^{29}Si and ^{30}Si can be explained in the same way as his results for $^{25,26}\text{Mg}$. That is, we assume that the energy of the (γ, n) giant resonance in ^{28}Si is also the energy of a component of the split giant resonance in $^{29,30}\text{Si}$. This has the value of 20.5 MeV. Katz also observed a peak in the cross section of the $^{29,30}\text{Si}(\gamma, n)$ reaction at 15 MeV. Unfortunately, it was not possible here to separate the effects of the two silicon isotopes being considered.

It is suggested that the 15 and 20.5 MeV peaks are the components of a split giant resonance as in the Danos-Okamoto effect. The agreement is not particularly good in this case.

The foregoing discussion of the experimental results for $^{29,30}\text{Si}$ has been entirely in terms of the properties of ^{29}Si . However, inclusion of the effect of ^{30}Si on the experimental results will not improve the results unless ^{30}Si is of prolate shape in contrast to the oblate shape specified for ^{28}Si and ^{29}Si by Bromley, Gove, and Litherland (1957). These authors indicate that ^{28}Si and ^{29}Si have the same value of the spheroidicity parameter δ . Further, a comparison of the spheroidicity parameter for the even-even nuclei ^{28}Si , ^{30}Si may be obtained from comparison of the energies of their first excited states (Bohr and Mottelson

1955). This comparison indicates that, if anything, ^{28}Si is more non-spherical than ^{30}Si . This observation is not in the direction to give better agreement with the experimental results, if we assume both ^{29}Si and ^{30}Si to be oblate spheroids. The assumption that ^{30}Si is a prolate spheroid would improve the agreement.

(d) ^{51}V ((γ, n) threshold, 11.08 MeV; (γ, p) threshold, 7.90 MeV)

Measurement of the intrinsic quadrupole moment of ^{51}V by the technique of microwave spectroscopy gave a result 0.6 ± 0.4 barn.

Coulomb excitation of the 0.32 MeV state in ^{51}V gives a value of $0.0056e^2 \times 10^{-48} \text{ cm}^4$ for the reduced transition probability $B(E2)$ of the excitation process. The ground state spin is $7/2^-$, and the first two excited states, of energy 0.32 and 0.93 MeV, probably have spins $5/2^-$ and $3/2^-$ respectively (Bunker and Starner 1955). The Coulomb excitation of the 0.32 MeV level is thus a $7/2^- \rightarrow 5/2^-$ transition, and has been verified to be an $E2-M1$ mixture. The relation between $B(E2)$ and Q_0 for this case is given by Bohr and Mottelson (1953) as

$$B(E2) = \frac{15}{16\pi} Q_0^2 \frac{K^2(I+1-K)(I+1+K)}{I(I+1)(2I+3)(I+2)}$$

for an $I+1 \rightarrow I$ transition. It is assumed that the three levels known form a rotational series having $K=3/2$. No other value of K will give a value of Q_0 within the wide limits of error of the spectroscopic moment. Using $K=3/2$, we obtain Q_0 as 0.51 barn, with an estimated error of ± 30 per cent.

Okamoto (1956) has successfully accounted for the width of the giant resonance in ^{51}V . An attempt is made here to estimate the detailed shape of the resonance by summing two Breit-Wigner shape resonances.

The data of Nathans and Halpern (1954) on the $^{51}\text{V}(\gamma, n)^{50}\text{V}$ reaction show some evidence for giant resonance splitting. With this in mind, an attempt was made to fit the experimental cross section values obtained by Nathans and Halpern with the sum of a pair of Breit-Wigner resonance curves of equal height. The lower energy peak was assumed to have width 2 MeV and the higher energy peak 4 MeV. The integrated cross sections of the two components thus satisfy one of the conditions of the Danos theory. The two widths were chosen as 2 and 4 MeV respectively because the same type of fit to the measured cross section for photon absorption in tantalum required peaks of these widths (Spicer, Thies, *et al.* 1958) and because these values for the widths evidently hold in the case of ^{25}Mg . The higher energy peak was fixed at 20 MeV and the position of the lower energy peak was adjusted until the best fit was found. The result of this fit is shown in Figure 1. The agreement with the experiment of Nathans and Halpern (1954) is excellent. The low energy peak is at 17 MeV according to this fit, and this leads, using Danos's theory, to a Q_0 of 0.67 ± 0.11 barn. The error quoted was estimated on the assumption that the energy separation has maximum inaccuracy of ± 0.5 MeV. This agrees with the value of Q_0 obtained from Coulomb excitation within the experimental errors of the two values.

(e) ^{63}Cu ((γ, n) threshold, 10.83 ± 0.02 MeV; (γ, p) threshold, 6.12 ± 0.01 MeV)

The nucleus ^{63}Cu was considered in the same way as ^{51}V . The spectroscopic quadrupole moment of ^{63}Cu has been measured, and is quoted as -0.16 barn ± 20 per cent. (Bleaney, Bowers, and Pryce 1955). This leads to $Q_0 = -0.80 \pm 0.16$ barn, and thus to the prediction that $\hbar\omega_b/\hbar\omega_a$ equals 0.86. In this case, the negative Q_0 means that the peak of smaller integrated cross section (at $\hbar\omega_a$) comes at the higher energy. A cross section shape for the $^{63}\text{Cu}(\gamma, n)$ reaction was computed on the same basis as that used to fit the

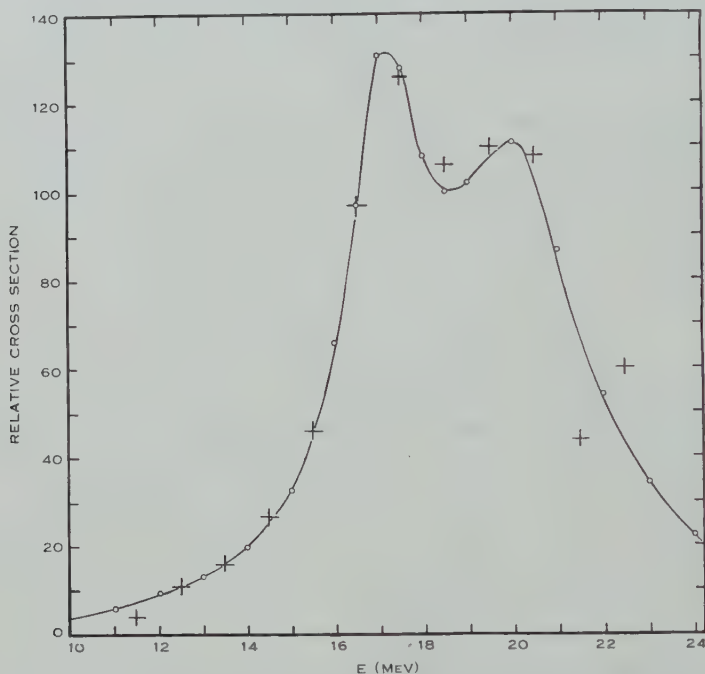


Fig. 1.—Fit of two Breit-Wigner resonance curves to the cross section for the $^{51}\text{V}(\gamma, n)^{50}\text{V}$ reaction. The resonance curves assumed had widths 2 and 4 MeV respectively, and were of equal height.

○ Predicted values, + experimental values.

vanadium cross section. The two peaks were assumed to have width 4 and 2 MeV respectively, and, for this fit, $\hbar\omega_a$ was assumed to have the value 19 MeV. This gave $\hbar\omega_b$ equal to 16.4 MeV. This computed cross section was compared in shape with the $^{63}\text{Cu}(\gamma, n)$ given by Katz and Cameron (1951). This experimental result was chosen because it used the smallest energy steps of any published cross section, namely 0.5 MeV. This fit is shown in Figure 2. Once again the agreement in shape is excellent.

(f) Other Nuclei

We now consider five other nuclei in the chosen atomic number range, in which this effect should be clearly observable. They are ^{23}Na , ^{27}Al , ^{39}K , ^{55}Mn , ^{59}Co . All save potassium are 100 per cent. isotopes, and ^{39}K is 93.38 per cent.

abundant in natural potassium. In all these cases the value of Q_0 is known as a result of either microwave spectroscopy or Coulomb excitation. The known data concerning these nuclei are exhibited in Table 2.

Using the Danos theory, the ratio of the energies of the two components of the split giant resonance was computed in each case.

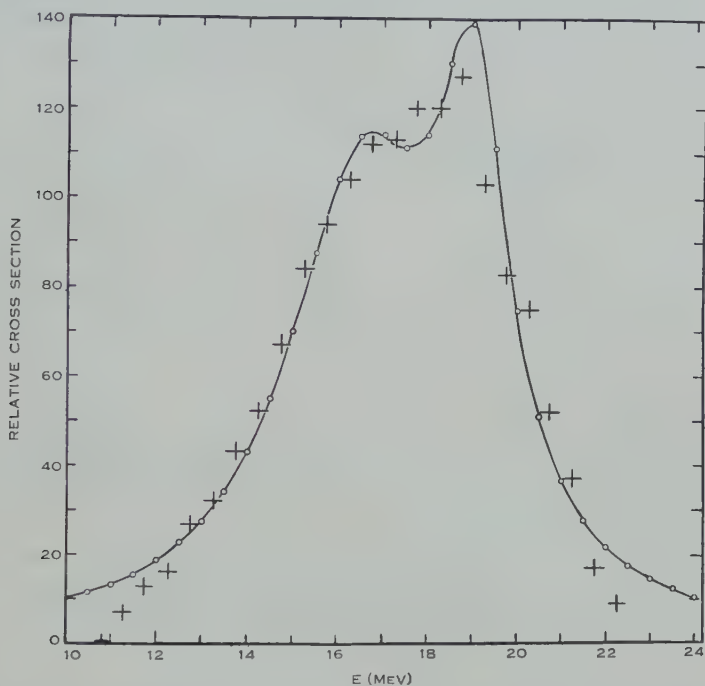


Fig. 2.—Comparison of $^{63}\text{Cu}(\gamma, n)^{62}\text{Cu}$ cross section with shape predicted using Danos's theory and the measured electric quadrupole moment of copper 63. The Breit-Wigner resonances assumed have widths 4 and 2 MeV respectively.

○ Predicted values, + experimental values.

To give an idea of the magnitude of the resonance splitting which may be expected, an energy value was assumed for the higher energy peak. The systematic study of (γ, n) reactions by Montalbetti, Katz, and Goldemberg (1953) was used as a guide in the choice of this value. However, the ratio of energies $\hbar\omega_b/\hbar\omega_a$ is the significant number in this result. The predictions are shown in Table 3.

For ^{23}Na the predicted energy of the low energy peak is at the (γ, n) threshold. In this case, the Danos-Okamoto effect should not be observable in the $^{23}\text{Na}(\gamma, n)$ reaction cross section, as the effect of competition from proton emission will be to keep the (γ, n) cross section low near threshold, and should therefore remove evidence of this low energy peak. The low energy peak should, however, be observable in the cross sections for the $^{23}\text{Na}(\gamma, p)$ and $^{23}\text{Na}(\gamma, \alpha)$ reactions. No

measurements have yet been made of these two cross sections. A prediction of the variation of the cross section for photon absorption with energy is shown in Figure 3. The peaks are assumed to have widths 2 and 4 MeV respectively, and to be of equal height.

TABLE 2
ECCENTRICITY AND PHOTONUCLEAR THRESHOLDS FOR ELEMENTS CONSIDERED

Nucleus	Q_0^* (barn)	ϵ	(γ, n) Threshold (MeV)	(γ, p) Threshold (MeV)
^{23}Na	$+0.50 \pm 0.05(\text{S})(\text{CE})^\dagger$	0.975	$12.41 \pm 0.02^\ddagger$	$8.25 \pm 0.03^\ddagger$
^{27}Al	$+0.44 \pm 0.01(\text{S})^\dagger$	0.652	$13.06 \pm 0.03^\ddagger$	$8.25 \pm 0.03^\ddagger$
^{39}K	$+0.70 \pm 0.17(\text{S})^\dagger$	0.56	$13.08 \pm 0.09^\ddagger$	$6.42 \pm 0.09^\ddagger$
^{55}Mn	$+1.54 \pm 0.8(\text{S})^\dagger$ $1.26 \pm 0.4^\S$	0.63	$10.15 \pm 0.25^\ddagger$	$8.16 \pm 0.20^\ddagger$
^{59}Co	$+1.1 \pm 0.4(\text{S})^\dagger$	0.47	$10.49 \pm 0.01 $	$7.41 \pm 0.01 $

* The method of measurement of Q_0 is indicated by the initial in parentheses after the value. (S) means a spectroscopically determined moment, and (CE) means that the moment is obtained from Coulomb excitation measurements.

† Value taken from Blin-Stoyle (1956).

‡ Thresholds computed from mass data given by Wapstra (1955).

§ This Coulomb excitation measurement was by Mark, McClelland, and Goodman (1955).

$||$ Thresholds given by Quisenberry, Scolman, and Nier (1956).

In the cases of ^{27}Al , ^{39}K , ^{55}Mn , and ^{59}Co , the predicted energy of the low energy peak is above the (γ, n) threshold, and therefore, in principle, the low energy peak should be observable. It is possible, however, that even in these

TABLE 3
PREDICTIONS OF THE RESONANCE ENERGY RATIO

Nucleus	$\hbar\omega_b/\hbar\omega_a$	Assumed $\hbar\omega_b$ (MeV)	Predicted $\hbar\omega_a$ (MeV)
^{23}Na	1.47 ± 0.06	18.3	12.5
^{27}Al	1.31 ± 0.01	19.7	15.1
^{39}K	1.26 ± 0.07	18.3	14.5
^{55}Mn	1.30 ± 0.08	20.5	15.8
^{59}Co	1.22 ± 0.08	18.0	14.8

cases the effect of competition from proton emission will prevent evidence for the low energy peak from appearing in the (γ, n) cross section for these nuclei. Again, the low energy peak should be present in the (γ, p) cross section. Halpern

and Mann (1951) describe measurements of the (γ, p) cross sections in ^{27}Al and ^{59}Co . In these two cases, the yield curve was not measured in sufficient detail to show the existence of a low energy peak. A programme of careful measurements of the low energy region of the (γ, n) cross sections and of the (γ, p) cross sections of the nuclei discussed in this section is under way in this laboratory, in an attempt to verify the predicted existence of the Danos-Okamoto effect in these nuclei.

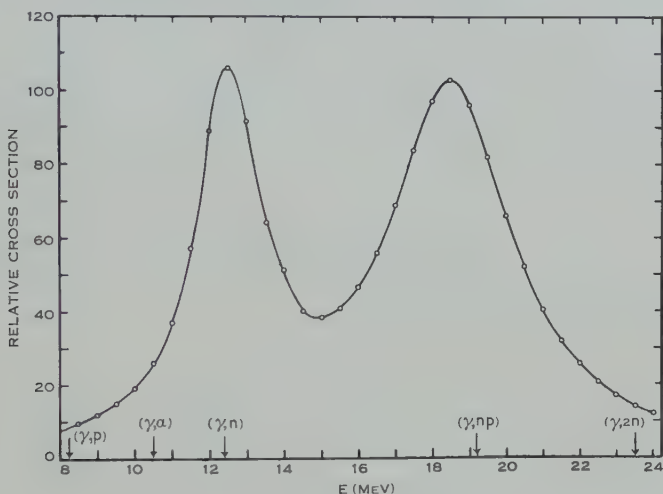


Fig. 3.—Predicted shape of the cross section for photon absorption in ^{23}Na . The various photoneuclear thresholds in ^{23}Na are indicated by arrows. Parameters of the resonances are :

$$\hbar\omega_a = 12.5 \text{ MeV}, \Gamma_a = 2; \hbar\omega_b = 18.5 \text{ MeV}, \Gamma_b = 4.$$

III. CONCLUSION

The predictions of the Danos-Okamoto theory on the splitting of the giant resonance into two component peaks are supported by experiments on the nuclei ^{19}F , ^{25}Mg , ^{29}Si , ^{51}V , and ^{63}Cu . The positions of the two peaks in the cross section for nuclear photon absorption are estimated for five other nuclei, but experimental data are either not good enough or are completely lacking in these cases. Therefore we cannot say whether or not the Danos-Okamoto effect occurs in these nuclei also. If it does, then the measurement of the cross section for nuclear photon absorption can be expected to give a measure of the deformation from spherical shape of that nucleus, and hence of the intrinsic electric quadrupole moment. In the region considered, the deformations are in many cases larger than those in the rare earth region, although the values of Q_0 are not large here.

One of the more striking results of this investigation concerns the widths of the components of the split giant resonance. The fit to the observed cross section for nuclear photon absorption in tantalum (Spicer, Thies, *et al.* 1958) required Breit-Wigner shapes of widths 2 and 4 MeV respectively. In the case of

^{25}Mg , where the peaks are well separated, these widths apply also. Further, the shapes of the giant resonance in ^{51}V and ^{63}Cu have been successfully fitted using components of this width. Thus it appears that these two widths pertain for the two components of the split giant resonance throughout the periodic table from $A=19$ up. If this is so, then the width of the giant resonance for photon absorption, in the case where the two components are not resolved, may be used to give a measure of the ground state nuclear deformation.

To do this, one would have to assume the correctness of the model used by Danos (1958) in his predictions of this effect. This model is the classical model of Steinwedel and Jensen (1950). One of the assumptions of this model which is not expected to be correct is the assumption of a rigid nuclear surface. However, the success of the Danos (1958) theory indicates that this assumption is at least a good approximation.

IV. ACKNOWLEDGMENT

The author wishes to thank Professor Sir Leslie Martin for his interest and encouragement throughout this work.

V. REFERENCES

- ALDER, K., BOHR, A., HUUS, T., MOTTELSON, B., and WINTHER, A. (1956).—*Rev. Mod. Phys.* **28** : 432.
- BLEANEY, B., BOWERS, K. D., and PRYCE, M. H. L. (1955).—*Proc. Phys. Soc. A* **66** : 410.
- BLIN-STOYLE, R. J. (1956).—*Rev. Mod. Phys.* **28** : 75.
- BOHR, A., and MOTTELSON, B. R. (1953).—*Math.-fys. Medd.* **27** : No. 16.
- BOHR, A., and MOTTELSON, B. R. (1955).—"Beta- and Gamma-ray Spectroscopy." (Ed. K. Siegbahn.) Ch. 17. (North Holland Publ.: Amsterdam.)
- BROMLEY, D. A., GOVE, H. E., and LITHERLAND, A. E. (1957).—*Canad. J. Phys.* **35** : 1057.
- BUNKER, M. E., and STARNER, J. W. (1955).—*Phys. Rev.* **97** : 1272.
- DANOS, M. (1958).—*Nuclear Phys.* **5** : 23.
- GOVE, H. E., BARTHOLOMEW, G. A., PAUL, E. B., and LITHERLAND, A. E. (1956).—*Nuclear Phys.* **2** : 132.
- HALPERN, J., and MANN, A. K. (1951).—*Phys. Rev.* **83** : 370.
- HEYDENBERG, N. P., and TEMMER, G. M. (1956).—*Annu. Rev. Nuclear Sci.* **6** : 77.
- HOFSTADTER, R. (1956).—*Rev. Mod. Phys.* **28** : 214.
- HORSLEY, R. J., HASLAM, R. N., and JOHNS, H. E. (1952).—*Phys. Rev.* **87** : 756.
- KATZ, L., and CAMERON, A. G. W. (1951).—*Canad. J. Phys.* **29** : 518.
- KATZ, L., HASLAM, R. N., GOLDBERG, J., and TAYLOR, J. G. (1954).—*Canad. J. Phys.* **32** : 580.
- LASICH, W. B., MUIRHEAD, E. G., and SHUTE, G. G. (1955).—*Aust. J. Phys.* **8** : 456.
- LITHERLAND, A. E., McMANUS, H., PAUL, E. B., BROMLEY, D. A., and GOVE, H. E. (1958).—*Canad. J. Phys.* **36** : 378.
- LITHERLAND, A. E., PAUL, E. B., BARTHOLOMEW, G. A., and GOVE, H. E. (1956).—*Phys. Rev.* **102** : 208.
- MAR, H., McCLELLAND, C., and GOODMAN, C. (1955).—*Phys. Rev.* **98** : 1245.
- MONTALBETTI, R., KATZ, L., and GOLDBERG, J. (1953).—*Phys. Rev.* **91** : 659.
- NATHANS, R., and HALPERN, J. (1954).—*Phys. Rev.* **93** : 437.
- NATHANS, R., and YERGIN, P. F. (1955).—*Phys. Rev.* **98** : 1295.
- NILSSON, S. G. (1955).—*Math.-fys. Medd.* **29** : No. 16.
- OKAMOTO, K. (1956).—*Progr. Theor. Phys., Osaka* **15** : 75.
- OKAMOTO, K. (1958).—*Phys. Rev.* **110** : 143.
- PAUL, E. B. (1957).—*Phil. Mag.* **2** : 311.

- QUISENBERY, K. S., SCOLMAN, T. T., and NIER, A. O. (1956).—*Phys. Rev.* **104**: 461.
- RAKAVY, G. (1957).—*Nuclear Phys.* **4**: 375.
- SHERB, R., LI, C. W., and CHRISTY, R. F. (1955).—*Phys. Rev.* **96**: 1258.
- SPICER, B. M., ALLUM, F. R., BAGLIN, J. E., and THIES, H. H. (1958).—*Aust. J. Phys.* **11**: 273.
- SPICER, B. M., THIES, H. H., BAGLIN, J. E., and ALLUM, F. R. (1958).—*Aust. J. Phys.* **11**: 298.
- STEINWEDEL, H., and JENSEN, J. H. D. (1950).—*Z. Naturf.* **5a**: 413.
- TOMS, M. E., and STEPHENS, W. E. (1951).—*Phys. Rev.* **82**: 709.
- WAPSTRA, A. H. (1955).—*Physica* **21**: 367.
- WEISS, M. S., and FULLER, E. G. (1958).—Proc. Washington Conference on Photonuclear Reactions (unpublished).

EXCITED STATES OF ^8Be FROM THE $^7\text{Li}(d,n)^8\text{Be}$ REACTION

By R. H. SPEAR*

[Manuscript received August 14, 1958]

Summary

The energy spectrum of neutrons from the reaction $^7\text{Li}(d,n)^8\text{Be}$ at 700 keV bombarding energy has been studied with nuclear emulsions at 0, 75, and 135°. Both resolution and statistics compare favourably with those of previous nuclear emulsion measurements of this spectrum. The results do not confirm the suggestions made by earlier workers that levels at 2.2, 4.1, 5.3, and 7.5 MeV in ^8Be participate in the reaction, and support the view that the only excited state below 9 MeV in ^8Be is the well-known broad level at 2.9 MeV.

I. INTRODUCTION

During the past decade there has been a good deal of controversy as to the level scheme of ^8Be in the region of excitation energy E_{ex} from 0 to 9 MeV. The experimental evidence, which has been amply reviewed elsewhere (see, for example, Titterton 1954; Holland *et al.* 1955; Nilson *et al.* 1958), may be broadly divided into two groups. Some experiments, especially those with good statistics, indicate that the well-known 2.9 MeV level is the only excited state in the region, whereas others, usually with relatively poor statistics, suggest a multiplicity of levels. In some respects the most convincing of the latter group of experiments are those using nuclear emulsions to study the energy spectrum of neutrons from the reaction $^7\text{Li}(d,n)^8\text{Be}$. The most recent of these is the work of Gibson and Prowse (1955); using 930 keV deuterons, they claimed evidence for the existence of levels at 2.2, 2.9, 4.1, and 5.3 MeV. The resolution achieved is apparently better than has been achieved in previous studies of the reaction, since the half width of the ground state group is between 400 and 450 keV. However, Trail and Johnson (1954*a*, 1954*b*) have used a neutron spectrometer consisting of a proton recoil telescope with thin polyethylene radiator to study the same reaction at a deuteron energy of 2.0 MeV. With much better statistics than the nuclear emulsion measurements they find no evidence for any states below 10 MeV other than the 2.9 MeV level. The resolution achieved, however, was relatively poor, the half width of the ground state group being 1100 keV. It is possible to argue that this lack of resolution precluded the observation of the fine structure suggested by the emulsion measurements; Gibson and Prowse themselves suggest that the groups were not observed because of a decrease in intensity relative to the 2.9 MeV group with increasing bombarding energy.

The aim of the present work, using nuclear emulsions to measure the energy spectrum of neutrons from the $^7\text{Li}(d,n)^8\text{Be}$ reaction, has been to achieve resolution comparable with the best of previous experiments, and hence to endeavour to confirm the existence of the suggested fine structure.

* Physics Department, University of Melbourne.

II. EXPERIMENTAL PROCEDURE

A 30 keV thick target of separated ^7Li was bombarded by 700 keV deuterons for an integrated exposure of 40,000 μC . The target thickness was estimated from the energy of scattered protons as measured with a 180° magnetic analyser. Neutrons were detected in pairs of 1-in. square 400μ Ilford C2 nuclear emulsion plates located at 15° intervals around the target, the camera and target arrangements being the same as those described previously (Bird and Spear 1957). The plates were processed using a conventional "temperature development" method. Tracks were measured and analysed using the procedures described in detail by Bird and Spear (1957).

III. RESULTS

Spectra have been obtained at 0° , 75° , and 135° . In each of the spectra illustrated, recoil protons were accepted for measurement if the recoil angle ψ was less than 25° , the dip angle in unprocessed emulsion, τ , was less than 12.4° , and the recoil proton range corresponded to E_{ex} less than 8.8 MeV. Except in Figure 5, no corrections have been applied for loss of tracks from the emulsion and variation with energy of the neutron-proton scattering cross section.

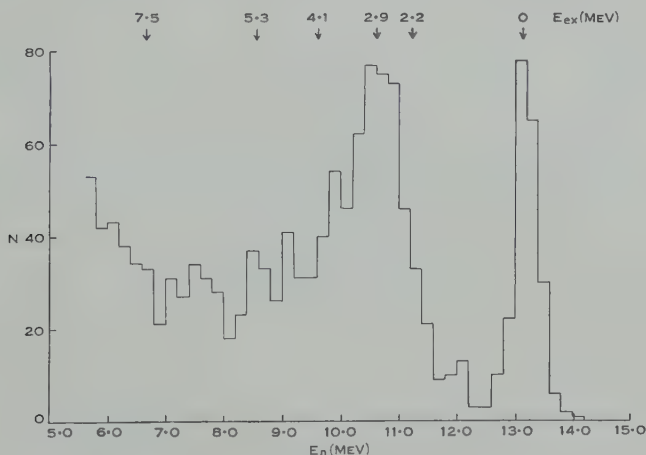
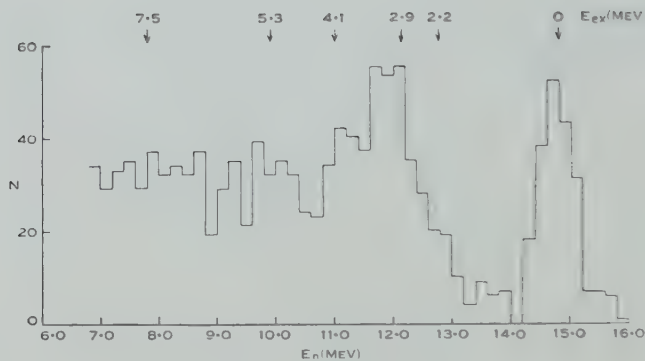
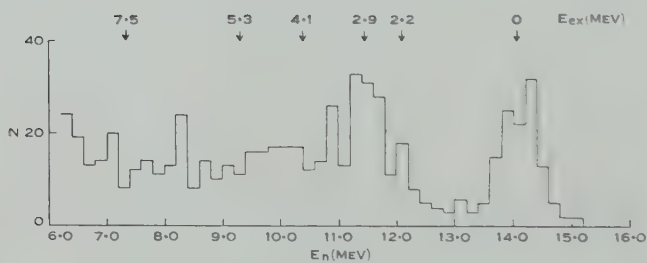
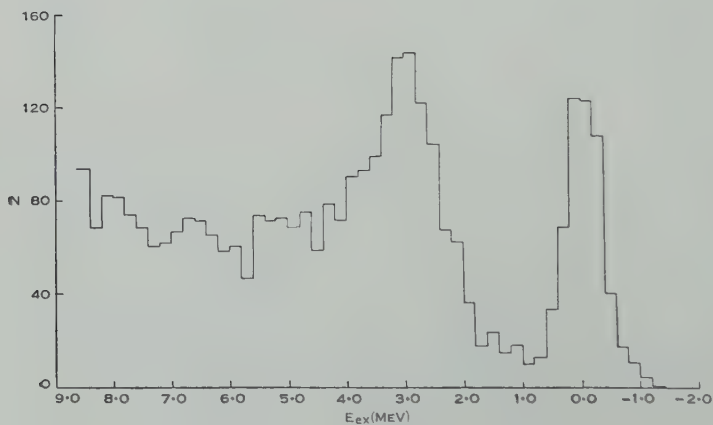


Fig. 1.—Spectrum at 135° .

The best resolution was obtained at 135° , the results for which are shown in Figure 1, where the number of tracks, N , per 200 keV interval is plotted against neutron energy E_n . The arrows indicate the positions at which groups would occur corresponding to excited states in ^8Be suggested by previous experiments. The resolution achieved is indicated by a half width of 440 keV for the ground state group, and the statistical accuracy is considerably better than that of Gibson and Prowse as determined from the statistical error bars on their published spectrum. The resolution may be improved further to 400 keV at the expense of statistical accuracy by limiting the acceptance angles ψ and τ more stringently; the shape of the spectrum is, however, essentially unaltered.

Fig. 2.—Spectrum at 0° .Fig. 3.—Spectrum at 75° .Fig. 4.—Combination of results at 0 , 75 , and 135° .

The results obtained at 0 and 75° are shown in Figures 2 and 3. The results for all three angles have been combined and plotted in Figure 4 in terms of ^8Be excitation energy. In Figure 5 the same results are shown after correcting for loss of tracks from the emulsion and variation with energy of the neutron-proton scattering cross section, statistical error bars being indicated at several points. The half width of the ground state group is approximately 700 keV. The total number of tracks represented in Figure 5 is 3300.

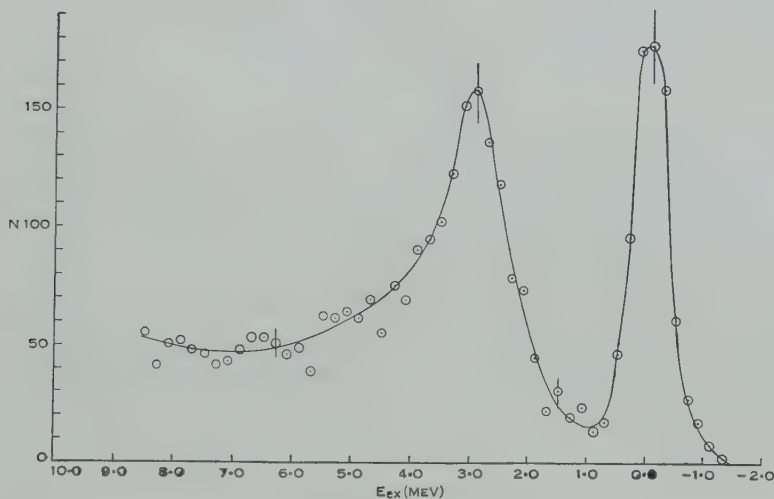


Fig. 5.—Combination of results at 0, 75° , and 135° after correcting for loss of tracks from the emulsion and variation with energy of the neutron-proton scattering cross section.

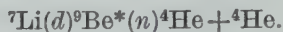
IV. DISCUSSION OF RESULTS

In none of the spectra illustrated is there any statistically significant indication of levels at 2.2, 4.1, 5.3, or 7.5 MeV. If such levels participate in this reaction in the way suggested from previous experiments, it is difficult to see why they should not have been observed in the present experiment, especially at 135° , where the resolution is similar to that achieved by Gibson and Prowse, and the statistical accuracy is considerably better.

From a measurement of the energies of α -particles emitted in coincidence with γ -rays in the reaction $^7\text{Li}(p, \gamma)^8\text{Be}(\alpha)^4\text{He}$, Inall and Boyle (1953) have suggested that transitions involving excited states at 4.1, 5.3, and 7.5 MeV in ^8Be occur with intensities of the order of 1 per cent. of the total radiation. Transitions of such low intensities would not have been detectable in the present work. However, LaVier, Hanna, and Gelinas (1956) used a magnetic spectrograph to study the energy spectrum of α -particles from the same reaction and found no evidence for the existence of such states, although transitions of 0.5 per cent. of the total intensity would have been detectable. Similar negative results have been obtained more recently by Meyer and Staub (1958) from magnetic analysis of the α -particle spectrum, both singly and in coincidence with γ -rays.

If there is no excited state below 10 MeV in ^8Be , other than the 2.9 MeV state, then some explanation is required for the continuum of neutrons for E_{ex} above 3 MeV. Three possibilities are suggested:

(i) A continuum of neutrons from the simultaneous three-particle break-up of ^9Be in the reaction



This reaction could provide neutrons of energies up to that of the ground state group from $^7\text{Li}(d,n)^8\text{Be}$.

(ii) A continuum of neutrons from the reaction



If ^5He is produced in its ground state, the maximum neutron energy is approximately 4.6 MeV at 0° . If excited states of ^5He are involved, more energetic neutrons may be obtained.

(iii) A contribution due to the overlapping of "tails" from groups corresponding to broad levels at 2.9 and 11 MeV. After correcting for the effects of experimental resolution, the half width obtained for the 2.9 MeV level is 1.6 ± 0.4 MeV. Values ranging from 0.8 to 2.0 MeV have been obtained from other reactions (Ajzenberg and Lauritsen, "Energy Levels of Light Nuclei, VI", personal communication). The evidence for the existence of a very broad level at about 11 MeV has been discussed by Ajzenberg and Lauritsen.

V. CONCLUSION

The results of this experiment do not confirm the deductions made from previous work that excited states at 2.2, 4.1, 5.3, and 7.5 MeV in ^8Be participate in the reaction $^7\text{Li}(d,n)^8\text{Be}$. They support the view that the only excited state in ^8Be below 9 MeV is the 2.9 MeV state.

VI. ACKNOWLEDGMENTS

The author is grateful to Professor Sir Leslie Martin, F.R.S., and to Dr. D. N. F. Dunbar for their helpful interest in this work. Thanks are also due to Dr. J. R. Bird for some useful criticisms, and to Mr. J. G. Rushbrooke, who assisted with some of the measurements.

VII. REFERENCES

- BIRD, J. R., and SPEAR, R. H. (1957).—*Aust. J. Phys.* **10**: 268.
 GIBSON, W. M., and PROWSE, D. J. (1955).—*Phil. Mag.* **46**: 807.
 HOLLAND, R. E., INGLIS, D. R., MALM, R. E., and MOORING, F. P. (1955).—*Phys. Rev.* **99**: 92.
 INALL, E. K., and BOYLE, A. J. F. (1953).—*Phil. Mag.* **44**: 1081.
 LAVIER, E. G., HANNA, S. S., and GELINAS, R. W. (1956).—*Phys. Rev.* **103**: 143.
 MEYER, V., and STAUB, H. H. (1958).—*Helv. Phys. Acta* **31**: 205.
 NILSON, R., JENTSCHKE, W. K., BRIGGS, G. R., KERMAN, R. O., and SNYDER, J. N. (1958).—*Phys. Rev.* **109**: 850.
 TITTERTON, E. W. (1954).—*Phys. Rev.* **94**: 206.
 TRAIL, C. C., and JOHNSON, C. H. (1954a).—*Phys. Rev.* **95**: 1363.
 TRAIL, C. C., and JOHNSON, C. H. (1954b).—*Bull. Amer. Phys. Soc.* **29** (7): 34.

OBSERVATIONS OF CHANGES IN THE PHOTOSPHERIC GRANULES

By R. J. BRAY* and R. E. LOUGHHEAD*

[*Manuscript received July 9, 1958*]

Summary

Using a sequence of 29 good-quality photographs extending over a period of 10 min, an attempt is made to detect changes in the brightness, size, and shape of individual granules. Of the 125 granules for which there is sufficient data, 57 per cent. show no detectable change during their observed periods of persistence, an additional 14 per cent. showing only minor changes of shape. Although there is some tendency among granules showing change for size increases to predominate over decreases, brightness increases and decreases occur with equal frequency. There is no correlation between the two types of change, nor is there any tendency for brightness or size variations to occur during any particular part of the life cycle. Observations with higher resolving power are required to elucidate the changes occurring during the formation and dissolution of individual granules.

In agreement with Macris, earlier estimates of the lifetimes of the granules are found to be too low. It seems likely that the true value for the most probable lifetime exceeds 7–8 min.

I. INTRODUCTION

Although the question of the mean lifetime of the photospheric granulation pattern has been investigated by a number of workers (for references see Macris 1953 ; Waldmeier 1955), no attempts have hitherto been made to detect changes in the brightness and size of the individual granules during their lifetimes, or to observe the details of their modes of formation and dissolution. Such observations, however, could give valuable information about the physical processes responsible for the granulation.

The reason for the lack of observations is the difficulty of photographing detail of the order of 1 sec of arc under conditions of day-time seeing. Moreover, while a few isolated photographs of moderate quality may suffice for estimating the mean lifetime of the pattern, a sequence of high-quality photographs extending over a period of at least 10 min is required if possible changes in the individual granules are to be detected.

In recent times the problem has been revolutionized by the application of time-lapse cinematography, and high-quality sequences of the granulation have been obtained by Lyot (in 1943) at the Pic-du-Midi (Macris 1953), by Blackwell, Dewhirst, and Dollfus (1957*a*, 1957*b*) from a manned balloon, and by Schwarzschild and his co-workers from a high-altitude unmanned balloon (Rogerson 1958). The results of the analyses of the balloon observations are awaited with interest.

* Division of Physics, C.S.I.R.O., University Grounds, Chippendale, N.S.W.

In the present paper an account is given of an attempt to detect changes in the brightness, size, and shape of the individual granules during the periods over which they were observed to persist. The photographs have been selected from a 10 min sequence obtained with a ground-level photoheliograph specifically designed for high-resolution photography. They are of sufficient quality to enable the granules to be classified according to brightness, size, and shape. It is found that the majority of the granules undergo no detectable change over the periods during which they were observed to persist. Although there is some tendency among granules showing change for size increases to predominate over decreases, brightness increases and decreases occur with equal frequency; there is no correlation between brightness and size variations. Finally, in agreement with Macris's results, the granules are found to remain identifiable over periods much greater than earlier estimates of their lifetimes would indicate.

II. OBSERVATIONS

The observations were made with the 5 in. photoheliograph (Loughhead and Burgess 1958) of the C.S.I.R.O. Division of Physics Solar Observatory. This instrument is designed to photograph any selected region of the solar disk on 35 mm film at 5-sec intervals. The diameter of the solar image is 20 cm and the effective wavelength 5400 Å. At this wavelength the theoretical limit of resolution is 0.8 sec of arc.

The present work is based on a sequence of the granulation in the neighbourhood of a spot group near the centre of the disk ($\cos \theta = 0.9$) photographed during the afternoon of February 4, 1958. The sequence contains 29 photographs of good but variable quality well distributed over a period of 10 min 21 sec. A portion of one of these photographs is reproduced in Plate 1. It will be noticed that the fundamental elements in the granulation pattern appear to be the bright ones, the dark material separating the bright areas being of no particular size or shape (cf. de Jager 1955; Leighton 1957). In this paper the term "granule" refers to the bright elements.

It is also apparent from Plate 1 that there is a considerable diversity in the brightness, size, and shape of individual granules. It might be argued that this diversity is illusory, being due simply to chance differences in the effect of atmospheric seeing on different regions of the photograph. However, while such differences do indeed occur, the fact that the diversity is real is readily shown by comparing the appearance of individual granules on several photographs taken within a minute, a period very much longer than the duration of any particular distortion due to seeing. It is then found that a given granule which on one of the photographs appears, for example, particularly large or perhaps elongated in a certain direction, in general appears similar on adjacent photographs (cf. Figs. 1-8). This fact, which is convincingly demonstrated when a large number of granules are examined on several photographs, shows that the photographs do to some extent reveal the true form and brightness of the granules and do not, as some writers have suggested, merely reproduce the instrumental profile of the telescope and atmosphere.

III. REDUCTION

The analysis of the enlargements of the 29 photographs (reproduced with normal contrast on such a scale that 1 mm corresponds to 0.6 sec of arc) was carried out in two stages. Firstly, 298 granules were selected from a "key" photograph occurring near the middle of the sequence. Each granule was given a number and its presence or absence on each of the remaining photographs was

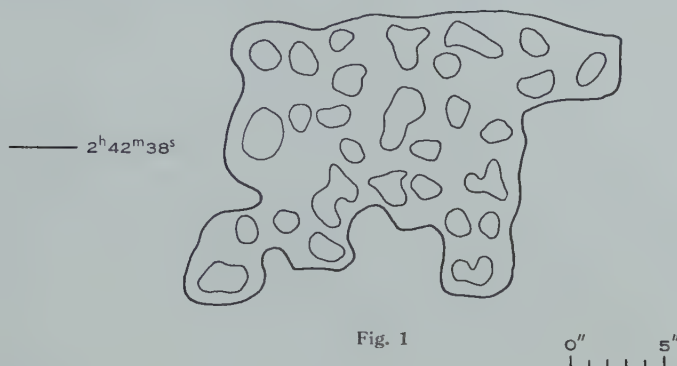


Fig. 1.—Granulation pattern at $2^h 42^m 38^s$. This figure is the "key" map for one of the eight selected regions (see text). In Figures 2–8 the pattern at $2^h 42^m 38^s$ (shown dotted) is compared with the pattern at earlier and later times. The pattern is easy to follow from map to map and it will be noticed that a number of granules persist throughout the sequence—a period of 9 min 6 sec. Despite apparent differences due to seeing, the sizes and shapes of individual granules are fairly well reproduced from figure to figure. For example, note how the elongation of the granule in the top right-hand corner persists from map to map.

recorded. Secondly, the granules were classified on the basis of their appearance at the time of the key photograph, and then an attempt was made to detect changes using good-quality photographs occurring before and after the time of the key.

A very good-quality photograph near the middle of the sequence was chosen as the key. On this photograph the 298 granules are distributed over eight

regions selected as being particularly free from the effects of seeing. In order to simplify the problem of identifying the same granules on successive photographs, it was decided to make maps of the granulation pattern in the eight regions on each of the 29 photographs (cf. Bray and Loughhead 1958). All maps were drawn independently, without reference to the key map or to one another. In tracing out the granules care was taken to draw outlines for only those bright features which were surrounded by dark material. No attempt was made to delineate the granules in any area seriously affected by seeing.

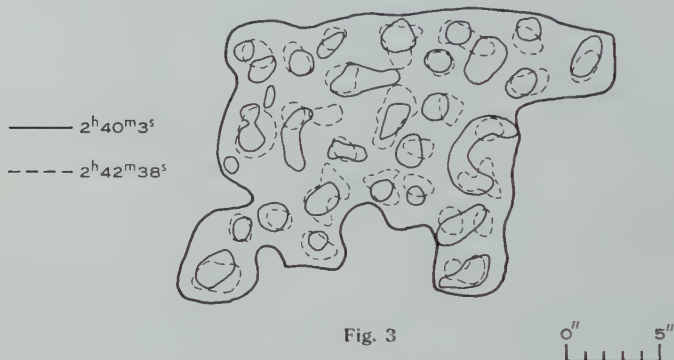


Fig. 3

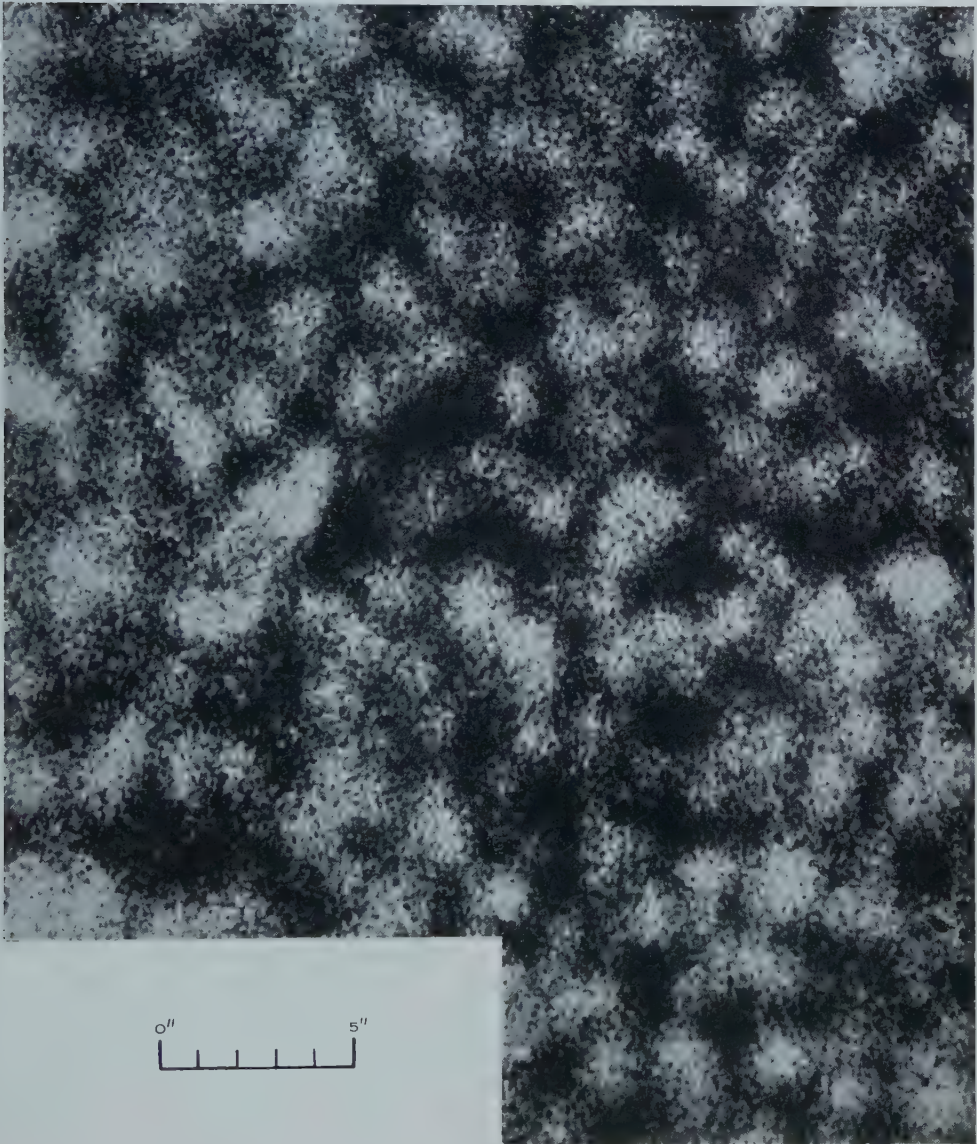


Fig. 4

The outlines of a few small adjacent sunspots were included on each map. The purpose of this was twofold, firstly as an aid in locating the eight selected regions on any given photograph, and secondly as an aid in comparing any given map with the key map. However, owing to small random displacements (~ 1 sec of arc) due to seeing, the final matching of the two maps could be made only by displacing one of them slightly until the best fit was obtained for the granulation pattern in the neighbourhood of the granules under study. The pattern is in fact quite easy to follow from photograph to photograph (cf. Figs. 1-8), so that the use of the auxiliary spots could have been avoided at the expense of additional labour.

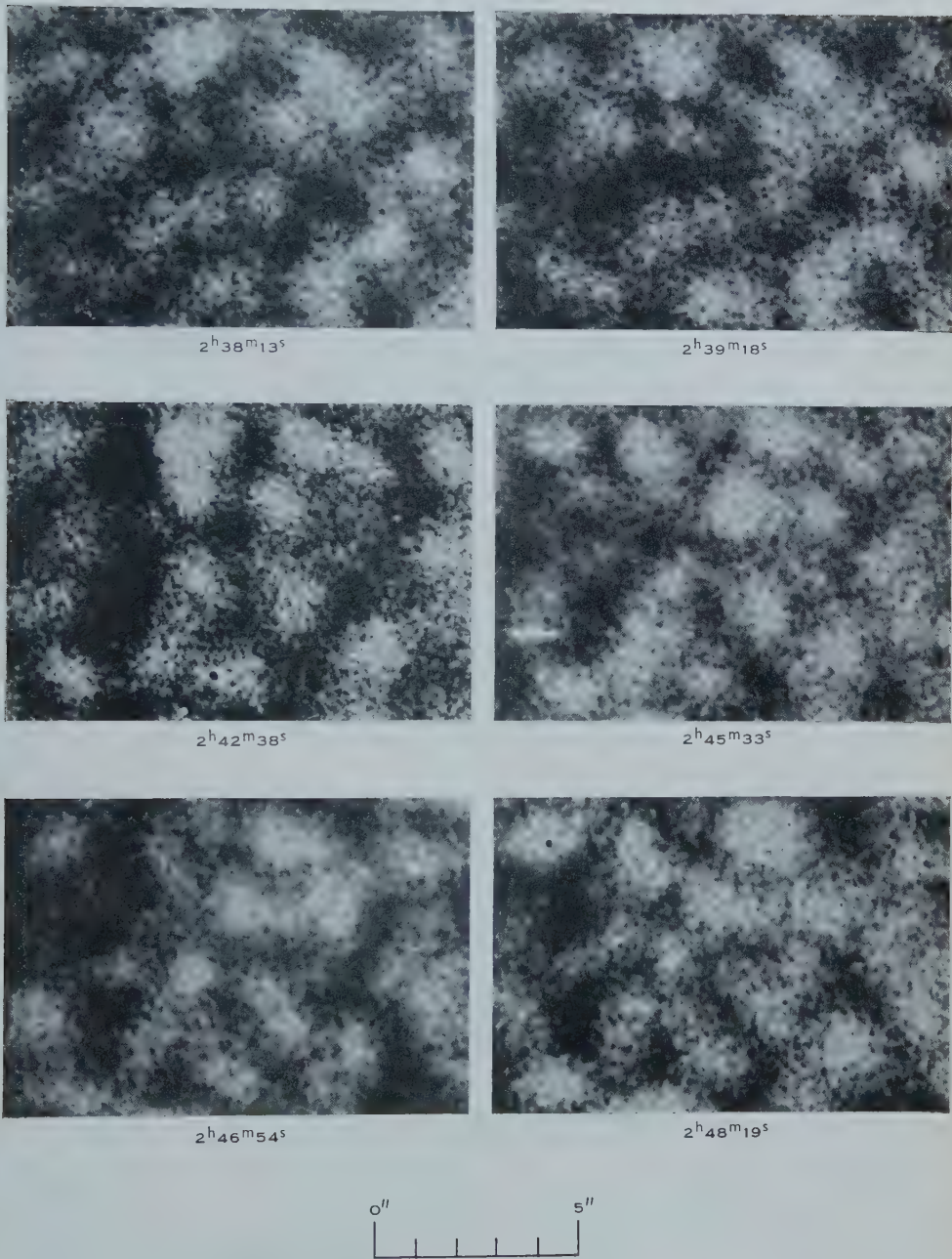
Each of the remaining 28 maps was compared in turn with the key map and the presence or absence at the corresponding time of each of the 298 granules

CHANGES IN THE PHOTOSPHERIC GRANULES



An enlargement from one of the 29 granulation photographs used in the present study. It will be noticed that the fundamental elements in the granulation pattern appear to be the bright ones, the dark material separating the bright areas being of no particular size or shape. There is a considerable diversity in the brightness, size, and shape of individual granules.

CHANGES IN THE PHOTOSPHERIC GRANULES



A sequence of photographs showing the stability of the granulation pattern over a period of 10 min 6 sec. Despite apparent changes due to seeing, several features can be followed without difficulty over almost the entire sequence.

was recorded. A positive identification for a given granule was allowed only if at least 50 per cent. of its area overlapped a granule on the key. In order to allow for the effect of seeing on the apparent size of a granule, or for any possible growth or decay, a positive identification was allowed when the areas were substantially different, provided the granules coincided in position.

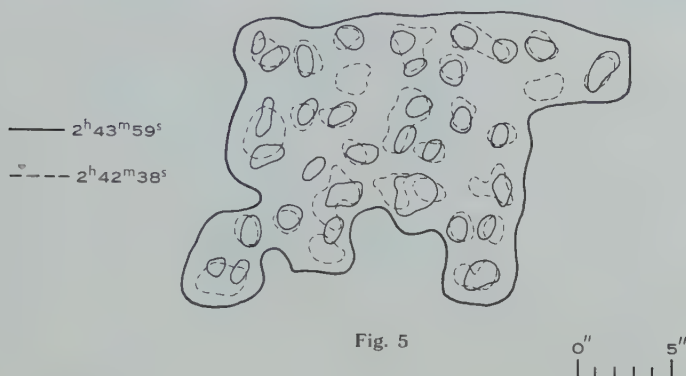


Fig. 5



Fig. 6

The above method of procedure is illustrated in Figures 1-8. The key map for one of the eight selected regions is shown in Figure 1 and in Figures 2-4 and 5-8 it is shown compared with maps corresponding to earlier and later times respectively. It will be noticed that a number of the granules persist throughout the sequence—a period of 9 min 6 sec. A discussion of the observed periods of persistence of the granules is given in Appendix I.

The small size of the granules, coupled with the limited resolving power, permits the granules to be described only in a rather crude and qualitative way; even this is only possible on less than one-half of the photographs. In fact, the description of the granules demands photographs much better than those required for mere identification. The classification of the granules was carried out on the basis of brightness, size, and shape. With regard to brightness the

categories were restricted to "bright", "medium", and "faint", and with regard to size, to "large", "average", and "small".*

Average-sized granules were classified as either "circular" or "elongated", whereas large granules were permitted an additional category, namely, "irregular". No shape classification was made for small granules. The direction of elongation, where applicable, was also recorded.

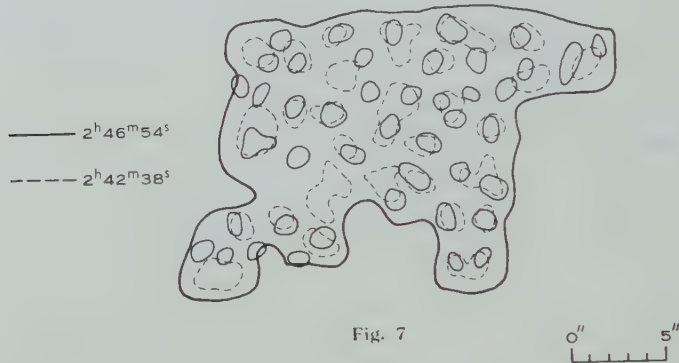


Fig. 7

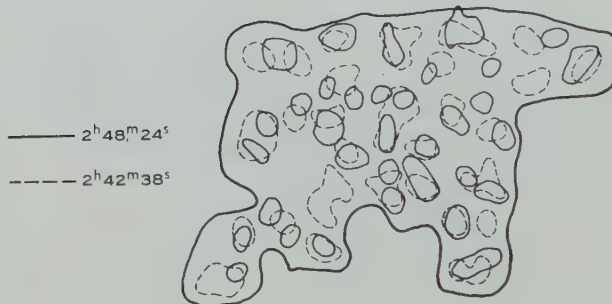


Fig. 8

The granules were initially classified according to their appearance at the time of the key photograph, using as supporting evidence three good photographs occurring within 35 sec. If a given granule either failed to appear on two or more of these photographs, or alternatively was too affected by seeing to be described, it was deleted. This procedure eliminated 158 of the original 298 granules. The remaining 140 were then described according to their appearance on each of a number of other selected photographs occurring before and after the key. This selection, which was varied with the region under study, contained

* The average measured diameters of the granules corresponding to these three classes are 1.8, 1.3, and 1.1 sec of arc respectively. However, these values should be regarded only as rough guides to the true dimensions since they depend to some extent on the contrast of the original negative and the positive print and on the combined instrumental profile of telescope and atmosphere.

on the average seven photographs, the remaining ones being of insufficient quality. On a given photograph, no attempt was made to describe a granule if it appeared to be appreciably affected by seeing.

IV. RESULTS

Table 1 gives the results of the classification of the granules at the time of the key, while Table 2 gives the results of the attempt to detect changes.

TABLE 1
CLASSIFICATION OF GRANULES*

Brightness			Size			
			Large	Average	Small	Total
Bright	12	23	2	37
Medium	23	50	19	92
Faint	1	6	4	11
Total	..		36	79	25	140

* The classification refers to the time of the key photograph.

TABLE 2
CHANGES IN GRANULES

Type of Change							No. of Granules
No change	71
Brightness	{ increase		12
	{ decrease		12
	{ increase and decrease*				4
Size	{ increase		16
	{ decrease		7
Change of shape	17
Total	125

* These granules showed both an increase and a decrease in brightness during the period of observation.

Considering large and average-sized granules together, Table 1 indicates that bright granules are about one-half as numerous as those of medium brightness and five times as numerous as faint granules. Particularly noteworthy is the almost complete absence of large, faint granules. With regard to small granules, the distribution appears to be quite different, bright granules being very rare and less numerous than faint ones. However, if the true size of these granules

is substantially smaller than the resolving limit of the telescope, the photographs would not reveal their true brightness, so this apparent distribution may well be erroneous. Observations with higher resolution may be expected to throw light on this question.

With regard to shape, 63 granules were found to be circular, 36 elongated, and 16 irregular. A considerable proportion are therefore non-circular (cf. Figs. 1-8). No preferred direction was found in the case of elongated granules.

Of the 125 granules for which there were sufficient data, 71 or 57 per cent. showed no detectable changes, an additional 14 per cent. showing only minor changes of shape (Table 2). The average period of observation was $6\frac{3}{4}$ min. Moreover, while there is some tendency among granules showing change for size increases to predominate over decreases, brightness increases and decreases occur with equal frequency. There is no correlation between the two types of change, nor was any tendency found for brightness or size variations to occur during any particular part of the life cycle.

The stability of the granulation pattern over a period of 10 min is illustrated in Plate 2. Although there are apparent differences from one photograph to another due to seeing, several features can be followed without difficulty over almost the entire sequence.

V. CONCLUSION

The observations show that the majority of the granules last considerably longer than earlier estimates would indicate (see Appendix I) and display little change during their observed periods of persistence. The observations refer only to granules in the neighbourhood of a sunspot group near the centre of the disk;* it is not known whether the results are applicable to granules in regions remote from spots.

The observations provide little information about the modes of formation and dissolution of individual granules. Only 26 cases of well-defined births or deaths were recorded; from these the impression was gained that in general a granule develops from a vague patch of diffuse, bright material which originates in a hitherto dark area. These diffuse patches are usually larger than the granules subsequently formed and are very difficult to distinguish from granules smeared by poor seeing. The dissolution of a granule appears to occur by the reverse process, although occasionally a granule loses its identity by coalescing with another granule. However, these remarks are only tentative and observations with much higher resolving power are required to make further progress in elucidating these aspects. Such observations may prove of decisive importance in choosing between rival theories of the mode of origin of the granulation.

VI. ACKNOWLEDGMENT

The authors wish to thank Mr. H. R. Gillett for processing the film and making the enlargements.

* The results refer, of course, to the true photospheric granulation and not to the "facular" granulation, since any faculae surrounding the spots would be transparent at the heliocentric angle of the group ($\cos \theta = 0.9$).

VII. REFERENCES

- BLACKWELL, D. E., DEWHIRST, D. W., and DOLLFUS, A. (1957a).—*Observatory* **77**: 20.
 BLACKWELL, D. E., DEWHIRST, D. W., and DOLLFUS, A. (1957b).—*Nature* **180**: 211.
 BRAY, R. J., and LOUGHHEAD, R. E. (1958).—*Aust. J. Phys.* **11**: 185.
 DE JAGER, C. (1955).—*Trans. Int. Astr. Un.* **9**: 727.
 LEIGHTON, R. B. (1957).—*Publ. Astr. Soc. Pacif.* **69**: 497.
 LOUGHHEAD, R. E., and BURGESS, V. R. (1958).—*Aust. J. Phys.* **11**: 35.
 MACRIS, C. (1953).—*Ann. Astrophys.* **16**: 19.
 ROGERSON, J. B. (1958).—*Sky & Telesc.* **17**: 112.
 WALDMEIER, M. (1955).—“*Ergebnisse und Probleme der Sonnenforschung.*” 2nd Ed. p. 92.
 (Geest u. Portig: Leipzig.)

APPENDIX I

Remarks on the Lifetimes of the Granules

Using a 22 min sequence of the granulation obtained by Lyot in 1943, Macris has derived values for the lifetimes of the granules considerably in excess of earlier estimates. He finds that the maximum of the lifetime distribution curve occurs for a value of 7–8 min. However, although individual lifetimes as high as 15–16 min were recorded,* the curve is weighted towards values less than 7 min, so that the mean value is 6–7 min.

The results of the present work are in broad agreement with those obtained by Macris and, in particular, indicate that earlier estimates are too low. Table 3

TABLE 3
OBSERVED TIMES OF PERSISTENCE OF GRANULES*

Time (min)	No. of Granules	Time (min)	No. of Granules
$0 < t < 1$	14	$6 < t < 7$	41
$1 < t < 2$	3	$7 < t < 8$	25
$2 < t < 3$	8	$8 < t < 9$	47
$3 < t < 4$	26	$9 < t < 10$	26
$4 < t < 5$	19	$t > 10$	23
$5 < t < 6$	17	Total ..	249

* These represent only *lower limits* to the true lifetimes (see text).

gives the periods during which the individual granules were observed to persist; it is based on measurements of 249 of the original 298 granules. (Forty-nine were rejected on the basis of their non-appearance on two or more of the three photographs occurring within 35 sec of the key.) The reliability of these values can be gauged from the fact that, for those granules which remained identifiable

* Macris finds that long lifetimes occur only for *clumps* of granules and concludes that individual granules have lifetimes of less than 8–9 min. This is not confirmed by the present work.

for 7 min or longer (nearly 50 per cent. of the total), positive identifications were recorded on the average for 20 photographs out of an average possible number of 22.

Owing to the relatively short duration of the sequence (10 min 21 sec), the starting and ending times of many of the granules fell outside the period of observation. Consequently the times given in Table 3 represent only *lower limits* to the true lifetimes. Therefore, although Table 3 appears to confirm Macris's value of 7-8 min for the most probable lifetime, it seems likely that a longer sequence (say 20 min or more) of equally good photographs would yield a higher value. Future observations alone can settle this question.

THE RADIO EMISSION FROM CENTAURUS-A AND FORNAX-A

By C. A. SHAIN*

[*Manuscript received June 13, 1958*]

Summary

Observations of the sources Centaurus-A and Fornax-A at 19.7 Mc/s, with a $1^{\circ}.4$ aerial beam, have been compared with observations at higher frequencies, in particular with Sheridan's 85.5 Mc/s observations. There are marked similarities between the sources, in their general appearances, linear dimensions, spectra, and powers radiated at radio frequencies.

More detailed discussion of the Centaurus-A observations strongly suggests that the extended component of this source has a spiral structure. It appears to be associated with the bright ellipsoidal part of NGC 5128, but possible association with the dark band cannot be excluded. One plausible interpretation of the data leads to an estimated age of the source of the order of 10^7 years.

I. INTRODUCTION

Although the number of radio sources now catalogued runs into thousands, only few of these can be identified, even tentatively, with optical objects, and the number reliably identified with external galaxies is even smaller. Two of this group, Centaurus-A and Fornax-A, associated with NGC 5128 and NGC 1316 respectively, are of particular interest, firstly in that the two galaxies, especially NGC 5128, are peculiar objects optically, and secondly because both sources have angular sizes which can be resolved with the present-day aerial systems of the Mills Cross type of radio telescope. It has already been shown by Sheridan (1958) that the Sydney 85.5 Mc/s Cross aerial, with its beamwidth of $0^{\circ}.8$, can be used effectively in the study of the brightness distributions across these two sources. Alongside the 85.5 Mc/s Cross at the Fleurs field station, near Sydney, N.S.W., is another Cross, working at a frequency of 19.7 Mc/s and with a beamwidth of $1^{\circ}.4$. This equipment has also been used to study Centaurus-A and Fornax-A.

The object of the present paper is first to describe the 19.7 Mc/s observations of these two sources and then to compare the results of these observations with those at other frequencies, especially Sheridan's observations at 85.5 Mc/s. Such a comparison of the available information about the two sources supports the hypothesis first advanced by de Vaucouleurs (1953) that the two associated galaxies have similar general physical characteristics. Finally, the great angular extent of the Centaurus source permits the study of its brightness distribution in some detail, and this leads to certain conclusions about the structure of the source.

* Division of Radiophysics, C.S.I.R.O., University Grounds, Chippendale, N.S.W.

II. THE 19.7 MC/S OBSERVATIONS

The main features of the aerial system and the method of recording have been described previously (Shain 1958). The radio telescope operates as a transit instrument, but aerial-switching gives quasi-simultaneous recording for five declinations separated by about 33 min of arc. The central declination can be changed by manual adjustment of the aerial, and the manual settings are usually arranged so that the northernmost of the five records for one setting should correspond to the southernmost record of the neighbouring setting, but in practice, owing to refraction in the ionosphere, the position in the sky to which the aerial is pointing at any time is different from the nominal position.

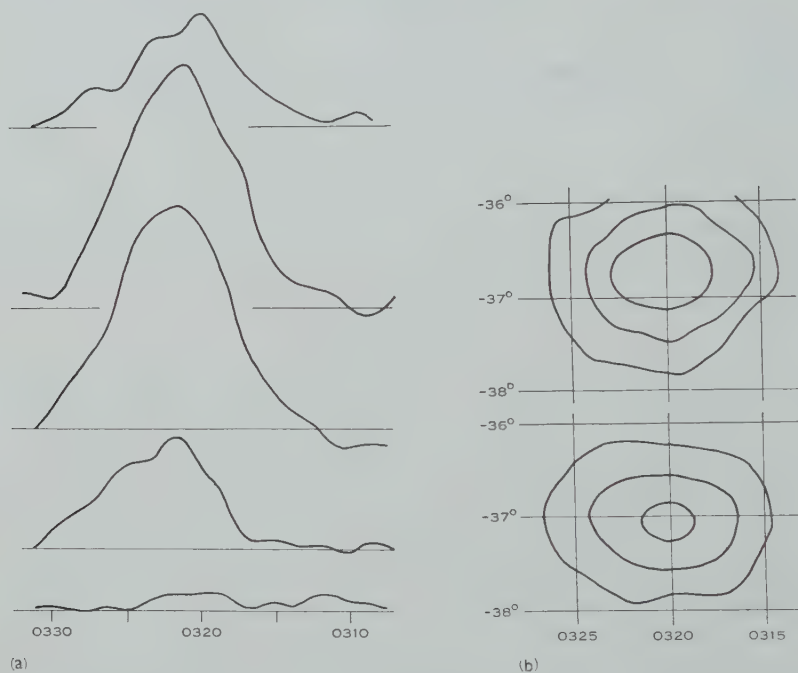


Fig. 1.—Observations of Fornax-A. (a) The five records at different declinations obtained on November 13, 1957. (b) Contour diagrams prepared from records such as those shown in (a). Upper diagram: November 13, 1957 (cf. (a)); lower diagram: November 7, 1957.

Both Centaurus-A and Fornax-A were observed only when they were within 10° of the zenith and all the observations were made between midnight and dawn. The F -region critical frequency was always less than 8 Mc/s, and, under these conditions, the expected refraction due to the curvature of the ionosphere should be small—less than about 3 min of arc (see, for example, Bailey 1948). However, the ionosphere is never simply stratified horizontally; at any time the electron density varies with both latitude and longitude, these variations producing apparent changes in the declination and Right Ascension respectively of extra-terrestrial sources. Smith (1952) has studied the refraction in Right Ascension at 81.5 Mc/s, but it has now been found that the refraction in declination is

equally important, and at 19.7 Mc/s the day-to-day variations of refraction in both coordinates may be $0^{\circ}.5$ or more.

An investigation of the relationship between the observed refraction and ionospheric conditions is being undertaken in this Laboratory, but so far there is insufficient information for the accurate forecasting from ionospheric sounding data of the refraction which might be expected. In any case, for the periods covered by the present observations, the ionospheric data from Australian stations was unfortunately incomplete. Because of the uncertain refraction, the records on different days were therefore made with the aerial beam pointing at directions in the sky which were not accurately known. Under these circumstances, the quasi-simultaneous recording of five separate declination scans was of very great value, making possible the empirical fitting together of the various records. As an illustration of the procedure adopted, the simpler case of the Fornax-A records will be outlined first.

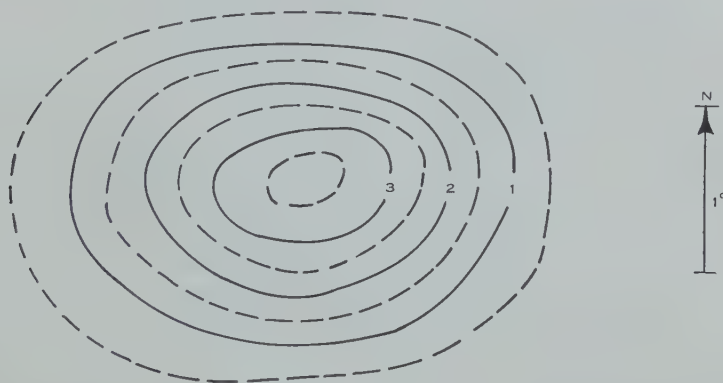


Fig. 2.—Isophotes of Fornax-A as observed at 19.7 Mc/s. The numbers on the contours give apparent brightness temperatures in units of $108,000^{\circ}\text{K}$.

In observing Fornax-A, the central position of the aerial beam was set close to the known declination of the source, and it was found that the five scans covered the source almost completely; tracings from one of the records are shown in Figure 1 (a). It was therefore possible to draw isophotes from each day's observations, and it was found that these isophotes had generally similar shapes (scintillations sometimes caused some distortion) but their apparent positions changed from day to day, as illustrated by Figure 1 (b). Four records were selected which apparently were free of scintillations and of any interfering signals, and the isophotes drawn from these records were superimposed in such a way that each set fitted the others as well as possible. Average brightness temperatures were then computed at intervals of 1 min in Right Ascension and 17 min of arc in declination and from these a composite set of isophotes was constructed. This is shown in Figure 2.

A similar process was adopted in reducing the Centaurus-A observations, but in this case observations with six different central positions of the aerial beam,

necessarily made on different days, were required to cover the great angular extent of the source. At least two records were taken for each beam position. Isophotes were again drawn for each day, and it was found that there was sufficient detail in them not only to match the records for one beam position, but also to match the isophotes for neighbouring beam positions with only very small

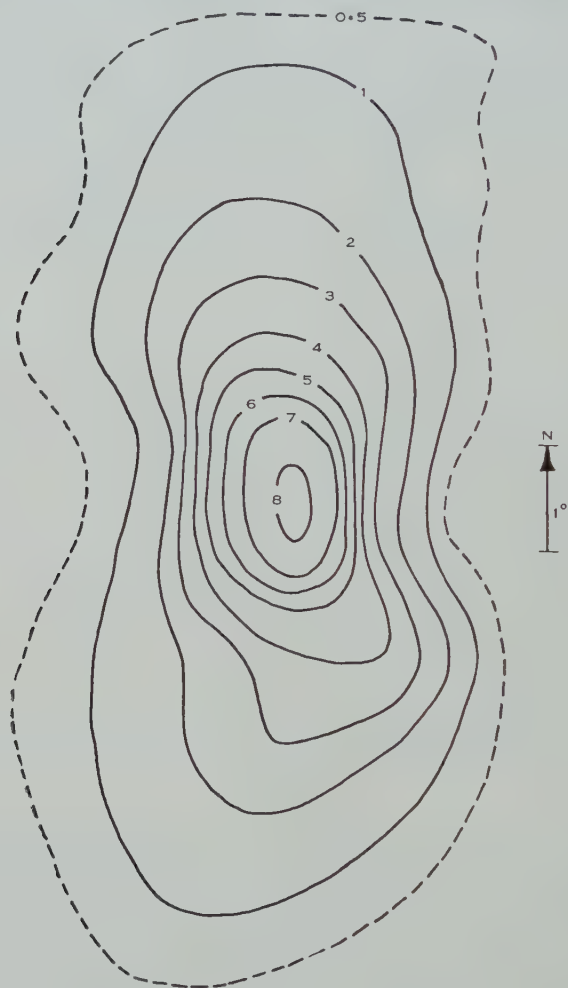


Fig. 3.—Isophotes of Centaurus-A as observed at 19.7 Mc/s. The numbers on the contours give apparent brightness temperatures in units of 100,000 °K.

uncertainty. In this way isophotes of the whole source were built up. The effects of the side lobes of the north-south aerial beam were removed in the manner described by Sheridan, and the effect of the background radiation, which changes only slowly in this region of the sky, was estimated and subtracted. The final result is shown in Figure 3, which depicts the isophotes of Centaurus-A as observed at 19.7 Mc/s.

The positions of the sources have not been indicated in Figures 2 and 3 because of the uncertain ionospheric refraction. The shapes of the 19.7 Mc/s isophotes are very similar to the shapes of the 85.5 Mc/s isophotes and it is assumed that the sources have the same positions at 19.7 Mc/s as at 85.5 Mc/s. Rough estimates of the refraction suggest that this assumption is correct. As regards intensities, relative values should be accurate to rather better than 10 per cent., but the uncertainty in the absolute scale is probably about 20 per cent.

Integration over the contours gives the following values for the flux densities of the sources :

Centaurus-A	..	$280 \times 10^{-24} \text{ W m}^{-2} (\text{c/s})^{-1}$, of which 11 per cent. is concentrated in an unresolved central source ;
Fornax-A	..	$43 \times 10^{-24} \text{ W m}^{-2} (\text{c/s})^{-1}$.

III. COMPARISON OF CENTAURUS-A AND FORNAX-A

(a) *Apparent Shape and Linear Dimensions*

A cursory inspection of Figure 3 shows that Centaurus-A is a very elongated source. In fact it may be objected that this extended source may be a chance superposition of a local galactic source on the "point" source which has already been identified with the dark obscuring band across the external galaxy NGC 5128. However, there are no other similar bright galactic features in this part of the sky, and the symmetry of the extended source about NGC 5128 is very strong evidence against the objection. If a rough allowance is made for the beamwidth of the aerial, the source extends for about $7\frac{1}{2}^\circ$ in the north-south direction and about $2\frac{1}{2}^\circ$ in the east-west direction. These dimensions are slightly larger than those observed at 85.5 Mc/s, and this is a real difference between the appearances of the source at the two frequencies, not just an effect of differing aerial resolution. The central "point" source is not resolved by either Cross.

Fornax-A is also an extended source. At 85.5 Mc/s it is only just resolved in the north-south direction, with an angular width of $0^\circ.5$ or less, but at both 85.5 and 19.7 Mc/s it is clearly resolved in the east-west direction, the angular lengths, allowing for aerial beamwidths, being about $0^\circ.8$ and $1^\circ.1$ respectively. Although it is not possible from the Cross results to say whether or not there is a strong central concentration as for Centaurus-A, Sheridan has pointed out that the complexity of the 85.5 Mc/s contours indicates some small-scale structure. It is possible that the greater angular extent at 19.7 Mc/s may be partly due to the greater contribution of an extended source to the total flux compared to that of a localized central source.

To get some idea of the linear dimensions of the two sources, it is necessary to estimate their respective distances, about which there is some uncertainty. For definiteness, we will take the estimates of de Vaucouleurs (1956), who gave the approximate distance of NGC 5128 as 7.5×10^5 pc, and of NGC 1316 as 5×10^6 pc. The distances can hardly be much larger than these since Baade and Minkowski (1954a) comment that, even if only at 5×10^5 pc, NGC 5128

must be one of the most luminous spheroidal galaxies. We then find that Centaurus-A extends over an area of about 100 by 30 kpc, and that the long dimension of Fornax-A is also about 100 kpc.

(b) *Spectra*

(i) *Centaurus-A*.—There have been many observations of this source, extending over a wide range of frequencies, but until the results of the Cross observations became available there was no reliable means of separating the contributions to the total flux density from the central source and the extended source, which has been called a “halo”, but which is now seen to be very elongated. It was noted above that at 19.7 Mc/s 11 per cent. of the flux is received from the unresolved central source. At 85.5 Mc/s the proportion is 23 per cent. (Sheridan 1958), and it is immediately clear that the two components must have different spectra.

In the light of the present observations, it is now possible to use earlier observations to estimate the spectrum of the two components separately. The sea-interferometer observations by Stanley and Slee (1950) at 60, 100, and 160 Mc/s and some of the observations by Mills (1952) at 101 Mc/s were made with the interferometer fringes so closely spaced that only the central source would be recorded. Observations at frequencies greater than 1000 Mc/s (Piddington and Minnett 1951; Haddock, Mayer, and Sloanaker 1954; Hagen, McClain, and Hepburn 1954) have been made using aerial beams ranging from $1^{\circ}.4$ to $0^{\circ}.5$. These aerials would have accepted only a small fraction of the radiation from the extended source, which, in any case, is comparatively weak at high frequencies. It follows that these very high frequency observations refer only to the central source, any small part of the extended source included in the aerial beams being faint and insignificant.

At 600 Mc/s Piddington and Trent (1956) give the integrated flux density from the whole source, and at 400 Mc/s McGee, Slee, and Stanley (1955) show contours from which the total flux density can be derived. At these two frequencies the contribution of the central source has now been estimated from the spectrum of this component, and this has been subtracted from the total flux density to give the flux density of the extended source. Shain and Higgins (1954) measured the total flux density at 18.3 Mc/s but no attempt has been made to separate the two components.

Figure 4 shows plots of the observed flux densities against frequency, and also lines corresponding to a variation of the flux density S with frequency f of the form $S = kf^{-n}$, where k and n are constants. In Figure 4 (a) the observed total flux density is plotted and the observations are fitted for a value of $n = 1.0$. In Figure 4 (b) the two components are separated and the values of n are 0.6 for the central source and 1.25 for the extended source. The difference in spectra of the two components is confirmed.

(ii) *Fornax-A*.—Because of its lower flux density, this source has been studied less extensively than Centaurus-A. Apart from the Cross observations, the total flux density has been measured by Shain and Higgins (1954) at 18.3 Mc/s, by Bolton *et al.* (1954) at 100 Mc/s, by McGee, Slee, and Stanley (1955) at 400 Mc/s,

and by Piddington and Trent (1956) at 600 Mc/s. The observed flux densities are plotted in Figure 4 (c), and, although there is some uncertainty, the value of n for Fornax-A must be close to 1.

(iii) *Ratio of Radio and Photographic Flux Densities*.—As extreme extensions of the spectra of the two sources, we may consider, for each source, the ratio of the flux density observed at a low radio frequency to the corresponding flux density observed in the photographic region of the spectrum. For this purpose we take the total flux densities observed at 85.5 Mc/s by Sheridan and the total photographic magnitudes listed by de Vaucouleurs (1956). These latter have been converted to flux densities using formulae given by Allen (1955, p. 174). We then find that the ratio of 85.5 Mc/s to photographic flux density for Centaurus-A is 1100, and for Fornax-A 1800.

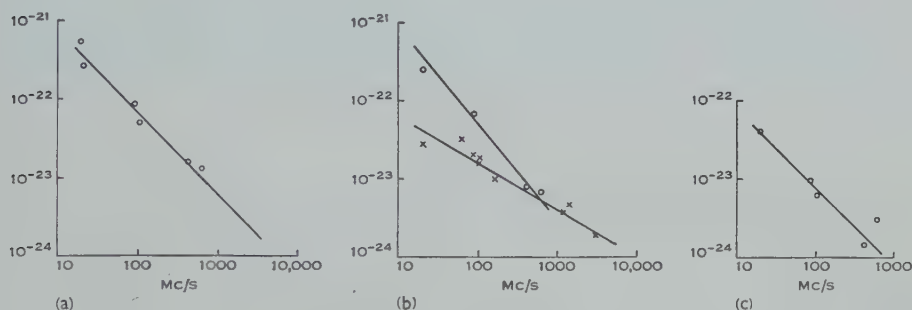


Fig. 4.—Spectra of Centaurus-A and Fornax-A. (a) Spectrum of the total flux density of Centaurus-A; the slope of the line is -1.0 . (b) Spectra of the two separate components of Centaurus-A. Circles: extended source, slope -1.25 ; crosses: central source, slope -0.6 . (c) Spectrum of the total flux density of Fornax-A, slope -1.0 .

(c) Power Radiated

As the values of n , describing the spectral variations of the sources, have been estimated, it is now possible to estimate the total power radiated in the radio-frequency range. In the absence of any observational evidence concerning cut-off frequencies, these were arbitrarily fixed, following Burbidge and Burbidge (1957), at 10 and 3000 Mc/s. Integration of the spectra then gives directly the radio-frequency flux received at the Earth, namely, for Centaurus-A (total) $3.9 \times 10^{-14} \text{ W m}^{-2}$, and for Fornax-A $4 \times 10^{-15} \text{ W m}^{-2}$. Taking the distances adopted above, we then find that the powers radiated by the sources are:

Centaurus-A (central source)	$1.0 \times 10^{39} \text{ erg/sec}$
(extended source)	$1.7 \times 10^{39} \text{ erg/sec}$
(total)	$2.7 \times 10^{39} \text{ erg/sec}$
Fornax-A	$1.2 \times 10^{40} \text{ erg/sec.}$

(d) NGC 5128 and NGC 1316

de Vaucouleurs' (1953) suggestion that these two galaxies appear to be similar has been criticized by Baade and Minkowski (1954b), who consider that the faint absorption patches of NGC 1316 are quite dissimilar to the heavy absorption band crossing NGC 5128. NGC 1316 is now included on Plate 11

of the Cape Photographic Atlas of Southern Galaxies and large patches of absorption are seen near the edge, even on this long-exposure photograph. Taken in conjunction with Baade and Minkowski's pictures, it appears quite likely that these absorption patches are confined to a band across the galaxy. It is clear from other observations by Minkowski (personal communication to Dr. B. Y. Mills), particularly of spectra, that there are definite differences between the two galaxies as observed optically. On the other hand, there does seem to be some similarity, and it appears that at present we cannot decide, on optical evidence alone, whether or not there is any fundamental similarity which would provide an explanation of their remarkable radio properties.

The Cape Atlas gives the extension of NGC 1316 as 6 min of arc, which would correspond to a length of 8700 pc at the adopted distance. Baade and Minkowski (1954a) can trace faint extensions of NGC 5128 over 30 min of arc, corresponding to a length of 6500 pc. Photoelectric recordings by de Vaucouleurs (1956) suggest that extremely faint extensions of this galaxy may be detected over about twice this distance, but no similar observations are available for NGC 1316.

TABLE 1
COMPARISON BETWEEN CENTAURUS-A AND FORNAX-A

	Centaurus-A	Fornax-A
Associated galaxy	NGC 5128	NGC 1316
Optical peculiarities	heavy absorption band	absorption patches in band across galaxy
Adopted distance (pc)	7.5×10^5	5×10^6
Absolute magnitude (pg)	-18	-19
Max. linear dimensions		
(optical) (pc)	6.5×10^3	8.7×10^3
(radio) (pc)	10^5	$\sim 10^5$
Ratio of length to breadth (radio)	3	> 2
Radio-frequency spectral index for whole source	-1.0	-1
Power radiated between 10 and 3000 Mc/s (erg/sec)	3×10^{39}	12×10^{39}
Ratio of 85.5 Mc/s and photographic flux densities	1.1×10^3	1.8×10^3

(e) *Summary of Evidence for the Similarity of Centaurus-A and Fornax-A*

The observational data on the two sources are collected together in Table 1. Some of the values in the table are based on the assumption that the distances adopted in Section III (a) are correct. As an example of the effect of changing the distance of one of the sources, if the distance of NGC 1316 were reduced to about 3×10^6 pc, so that its absolute magnitude was the same as that of NGC 5128, its linear dimensions would be 5.5×10^3 pc (optical) and 6×10^4 pc (radio), and the radiated power would be about 5×10^{39} erg/sec.

All the data of Table 1 are consistent with the hypothesis that, although there may be differences in detail, Centaurus-A and Fornax-A are similar objects

physically. Taken together, they would appear to offer some support to this hypothesis, but further optical evidence and the study of the relationship between these and other "radio galaxies", such as M 87 (Virgo-A), (see Section IV), are required before it can be considered as definitely established. If it is accepted, the interpretation of NGC 5128 as a pair of colliding galaxies (Baade and Minkowski 1954a) would then appear to be ruled out, as such a description could hardly be applied to NGC 1316.

IV. FURTHER DISCUSSION OF CENTAURUS-A

Because it has such a large angular size, the extended component of Centaurus-A can be studied in some detail with the Cross equipment, and we shall now use the 19.7 and 85.5 Mc/s observations to see what information can be obtained about the structure of this extended source and its association with NGC 5128. It will be assumed in the subsequent discussion that the radiation originates in the acceleration of relativistic electrons in magnetic fields. On this assumption, and using the experimental evidence available at the time,

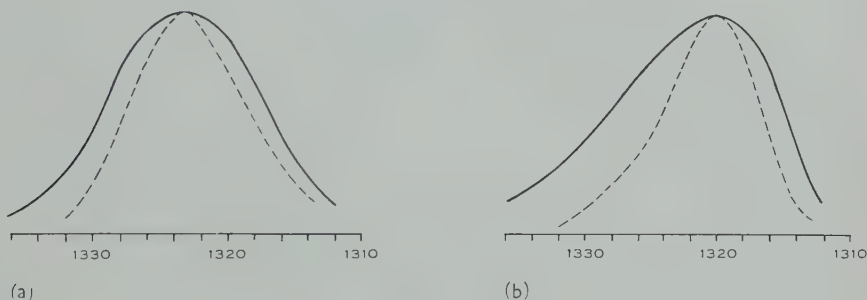


Fig. 5.—Sections across Centaurus-A at two declinations: (a) -41° , (b) $-44^\circ.5$. The full and dashed curves are 19.7 and 85.5 Mc/s observations respectively.

Burbidge and Burbidge (1957) have already calculated the total particle and magnetic field energies required to produce the observed emission. These energy estimates depend on the square of the adopted distance to the source, but they are not very sensitive to the values of the spectral indices of the source components. As Burbidge and Burbidge adopted a very large value of the distance, it is thought that the energies they deduced are rather too high, by about one order of magnitude, but otherwise the new data do not appear to affect their general argument.

In Section III (a) it was stated that the extended source appears to be wider at 19.7 than at 85.5 Mc/s. The source lies almost along an hour circle, so Figure 5 shows sections across the source at declinations -41° and $-44^\circ.5$, the scales being adjusted so that the maxima coincide.* From this figure we find for the angular width to half power at the two declinations: $1^\circ.9$ and $1^\circ.6$ at 85.5 Mc/s; $2^\circ.6$ and $2^\circ.8$ at 19.7 Mc/s. These are all about twice the aerial

* An approximate allowance for the general background radiation has been removed from the 85.5 Mc/s data using information kindly supplied by Mr. Sheridan. Any small uncertainties in doing this do not affect Figure 5.

beamwidths, so that the corrections for the width of the aerial beams are small (about $0^{\circ}\cdot 2$ at $85\cdot 5$ Mc/s and $0^{\circ}\cdot 4$ at $19\cdot 7$ Mc/s). In Figure 5 it is also seen that the positions of the maxima are the same at the two frequencies but the maximum for one declination is displaced from the maximum for the other declination.



Fig. 6.—Showing the spiral shape of the extended component of Centaurus-A. The heavy dashed line shows the ridge line of the source through the circles ($19\cdot 7$ Mc/s observations) and crosses ($85\cdot 5$ Mc/s observations).

Sheridan has already remarked that the extended source has a shape suggestive of an open spiral. This shape is not so marked at $19\cdot 7$ Mc/s (see Fig. 3); it would be masked by the faint but more extensive regions observed at this frequency. Nevertheless, there is a suggestion of a spiral shape and to bring this out Figure 6 has been drawn. In this figure, the “ridge line” of the extended source has been drawn by locating the points on the contours at which the curvature is a maximum—a simpler procedure than the drawing of sections

as in Figure 5, but giving the same result. The region very close to the central concentration was not considered. Both 85.5 and 19.7 Mc/s data agree well, and the heavy dashed line drawn through the points shows a definite spiral shape.

The spiral is asymmetrical about the centre, but optical photometric studies of NGC 5128 (Evans 1949; de Vaucouleurs and Sheridan 1957) suggest that the nebula is being viewed from the south. (The dark band appears slightly to the north of the centre of the bright nebula, and the southern part is brighter than the northern.)

The plausible model of the extended source is then a spiral, slightly more open than that indicated by Figure 6, with the south-following end nearer the observer and with the plane of the spiral inclined to the line of sight. Asymmetry is produced because the great extent of the source means that the two arms are at different distances and are viewed at different angles. On any model, the greater width at 19.7 Mc/s is readily explained by a weakening of the magnetic fields near the edge of the system and the gradual loss of energy by the highest energy electrons as they diffuse outwards from the centre.

We may note that the apparent position angle of the central part of the spiral is within 10° of the major axis of the main body of the nebula as shown in a long-exposure photograph reproduced by Baade and Minkowski (1954a). This might suggest that the extended source is associated with the bright part of the nebula and the central concentration with the dark band (Mills 1953). Alternatively, it is possible that the ends of the dark band curl around into this direction; the whole source would then have a barred spiral structure with a bright concentration in the very small bar. It is not possible to study such detail near the centre with the Cross equipment, but that some complication is present in that region is clear from Mills's (1953) interferometer observations. The whole question is complicated by the fact that the dark band and the major axis of the nebula are very nearly at right angles. This may well mean that the shape of the nebula (and of the extended source) is somehow intimately connected with the appearance of the dark band.

Any discussion of the origin of the source's spiral shape must be largely speculative until some decision can be made concerning the way in which the electrons (or positrons) responsible for the radio emission reach their high energies. Burbidge and Burbidge (1957) have already shown that the electrons must have been produced at high energies, since acceleration processes are too inefficient. An estimate of the number density of the particles actually responsible for the emission can be obtained by using the theory outlined by Ginzburg (1956). We then find that if the magnetic field is 10^{-5} G the particle density is $4 \times 10^{-13} \text{ cm}^{-3}$, whilst for a magnetic field of 10^{-6} G it is $7 \times 10^{-11} \text{ cm}^{-3}$. The total number of particles would be about 10^{57} or 10^{59} respectively, and even if there were an equal number of particles of protonic mass the total rest mass would be only 1 or 100 solar masses. Of course, the mass of the extended source may well be much higher than this. For example, Burbidge and Burbidge considered in detail the production of the high-energy electrons by collisions of a primary high-energy proton flux with static gas and dust in the extended source,

and their assumed density of about 10^{-26} g/cm³ (about 0.01 proton/cm³) leads to a mass of about 10^{10} solar masses. This would be an appreciable part of the mass of the NGC 5128 system and the general theory of galactic structure would have to be considered.

On the other hand it is possible that the radiating particles are themselves ejected at high energy from the galaxy, so that the mass of the source is extremely small. Then the spiral shape is a consequence of the rotation of the nucleus through roughly one-quarter of a revolution since the start of the ejection. Using Mayall's (1948) data on rotation periods of galaxies, the age of the extended source would be about 10^7 years, and the arms would expand at the rate of about 3000 km/sec. In this regard, it is noteworthy that Baade and Minkowski (1954a) quote spectral observations, by Humason, of M 87 (identified with the radio source Virgo-A) which suggest that the matter in the jet emerging from the nucleus of this galaxy is being ejected at speeds of the order of 500 km/sec. They estimate the age of the jet to be about 10^6 years. These figures would seem to favour Burbidge and Burbidge's suggestion that M 87 represents an early stage in the development of a system such as NGC 5128. This suggestion ties in with the hypothesis of Ambartsumian (1958) that NGC 5128 is actually a galaxy in the process of division.

It would be very interesting if a detailed spectroscopic study of NGC 5128 gave evidence of a rotation of the system consistent with this interpretation of the radio data observations. In fact, such a study would seem to be very desirable just to supplement the meagre optical information at present available.

V. ACKNOWLEDGMENTS

The author is indebted to a number of his colleagues for their cooperation in the present work. Thanks are particularly due to Mr. L. F. Clague for his help with the equipment, and to Miss J. Self for her help with the reduction of the records.

VI. REFERENCES

- ALLEN, C. W. (1955).—"Astrophysical Quantities." (Athlone Press: London.)
- AMBARTSUMIAN, V. A. (1958).—"The Large-scale Structure of the Galactic System." I.A.U. Symposium No. 5, pp. 4-6. (Cambridge Univ. Press.)
- BAADE, W., and MINKOWSKI, R. (1954a).—*Astrophys. J.* **119**: 215-31.
- BAADE, W., and MINKOWSKI, R. (1954b).—*Observatory* **74**: 130-1.
- BAILEY, D. K. (1948).—*Terr. Mag. Atmos. Elec.* **53**: 41-50.
- BOLTON, J. G., WESTFOLD, K. C., STANLEY, G. J., and SLEE, O. B. (1954).—*Aust. J. Phys.* **7**: 96-109.
- BURBIDGE, G. R., and BURBIDGE, E. M. (1957).—*Astrophys. J.* **125**: 1-8.
- EVANS, D. S. (1949).—*Mon. Not. R. Astr. Soc.* **109**: 94-102.
- GINZBURG, V. L. (1956).—*Nuovo Cim.* **3** (Suppl. No. 1): 38-48.
- HADDOCK, F. T., MAYER, C. H., and SLOANAKER, R. M. (1954).—*Astrophys. J.* **119**: 456-9.
- HAGEN, J. P., McCLAIN, E. F., and HEPBURN, N. (1954).—*Proc. Inst. Radio Engrs., N.Y.* **42**: 1811.
- MCGEE, R. X., SLEE, O. B., and STANLEY, G. J. (1955).—*Aust. J. Phys.* **8**: 347-67.
- MAYALL, N. U. (1948).—*Sky & Telesc.* **8**: 3-5, 17.
- MILLS, B. Y. (1952).—*Aust. J. Sci. Res. A* **5**: 266-87.
- MILLS, B. Y. (1953).—*Aust. J. Phys.* **6**: 452-70.

- PIDDINGTON, J. H., and MINNETT, H. C. (1951).—*Aust. J. Sci. Res. A* **4**: 459–75.
- PIDDINGTON, J. H., and TRENT, G. (1956).—*Aust. J. Phys.* **9**: 74–83.
- SHAIN, C. A. (1958).—*Proc. Inst. Radio Engrs., N.Y.* **46**: 85–8.
- SHAIN, C. A., and HIGGINS, C. S. (1954).—*Aust. J. Phys.* **7**: 130–49.
- SHERIDAN, K. V. (1958).—*Aust. J. Phys.* **11**: 400–8.
- SMITH, F. G. (1952).—*J. Atmos. Terr. Phys.* **2**: 350–5.
- STANLEY, G. J., and SLEE, O. B. (1950).—*Aust. J. Sci. Res. A* **3**: 234–50.
- DE VAUCOULEURS, G. (1953).—*Observatory* **73**: 252–4.
- DE VAUCOULEURS, G. (1956).—*Occ. Notes R. Astr. Soc.* **3**: 118–42.
- DE VAUCOULEURS, G., and SHERIDAN, K. V. (1957).—“Radio Astronomy.” I.A.U. Symposium No. 4, p. 169. (Cambridge Univ. Press.)

A PENCIL-BEAM SURVEY OF THE GALACTIC PLANE AT 3.5 M

By E. R. HILL,* O. B. SLEE,* and B. Y. MILLS*

[Manuscript received June 24, 1958]

Summary

A survey has been made of the galactic plane region from $l=223^\circ$ through the galactic centre to $l=13^\circ$ between $b=+4^\circ$ and -6° using the 3.5 m wavelength cross-type aerial (beamwidth 50 min of arc) near Sydney. Contour diagrams of brightness temperature have been prepared. The preparation of contours is described in detail, and a detailed discussion of the accuracy of the temperatures is given.

I. INTRODUCTION

In an effort to understand the origin of radio-frequency radiation emitted by the Galaxy, numerous surveys of the sky have been carried out in the past at various frequencies. Until 1954, however, there were no aerials operating at metre wavelengths capable of matching the angular resolution obtainable at centimetre wavelengths. The great physical size required of conventional-type aerials made it impracticable to obtain a steerable pencil beam for work at long wavelengths.

That there were features of interest in the distribution of radiation at long wavelengths, especially near the galactic plane, was made quite evident by the interferometer investigations by Scheuer and Ryle (1953) at 1.4 and 3.7 m, and by Bolton *et al.* (1954) at 3 m. The former were able to show the presence of intense radiation originating within a degree or so of the galactic plane, but no decision could be made as to how this radiation was distributed along the plane since the resolution of the aerial in this direction was low. The work of Bolton *et al.* clarified the situation somewhat in that it showed that there were regions of emission closely confined to, and elongated along, the galactic plane. These features were irretrievably lost in the lower resolution surveys.

The investigations noted above indicated, furthermore, that the most satisfactory method of examining the expected complicated distribution of radiation in the vicinity of the galactic plane is by the use of a pencil beam of width preferably less than 1° . That such resolution could be obtained at metre wavelengths by means of a cross-type aerial system was shown by Mills and Little (1953).

The purpose of the present communication is to describe the construction of brightness contours in a strip along a considerable length of the galactic plane using observational material obtained with the 3.5 m cross-type aerial situated at Fleurs, near Sydney. This aerial, which has a beamwidth of 50 min of arc, has been described together with its associated equipment by Mills, Little, *et al.*

* Division of Radiophysics, C.S.I.R.O., University Grounds, Chippendale, N.S.W.

(1958). As this paper will be referred to frequently in the course of subsequent discussion, it will be called paper I.

With such an instrument, detailed comparison between high and low frequency surveys, which had hitherto been restricted to the less complex regions away from the galactic plane and cooler regions near the anticentre, may now be extended right into the galactic plane. A brief preliminary discussion of the more interesting astronomical implications of the present high resolution survey has been given by Mills, Hill, and Slee (1958). More detailed analyses of these and other aspects are now in progress.

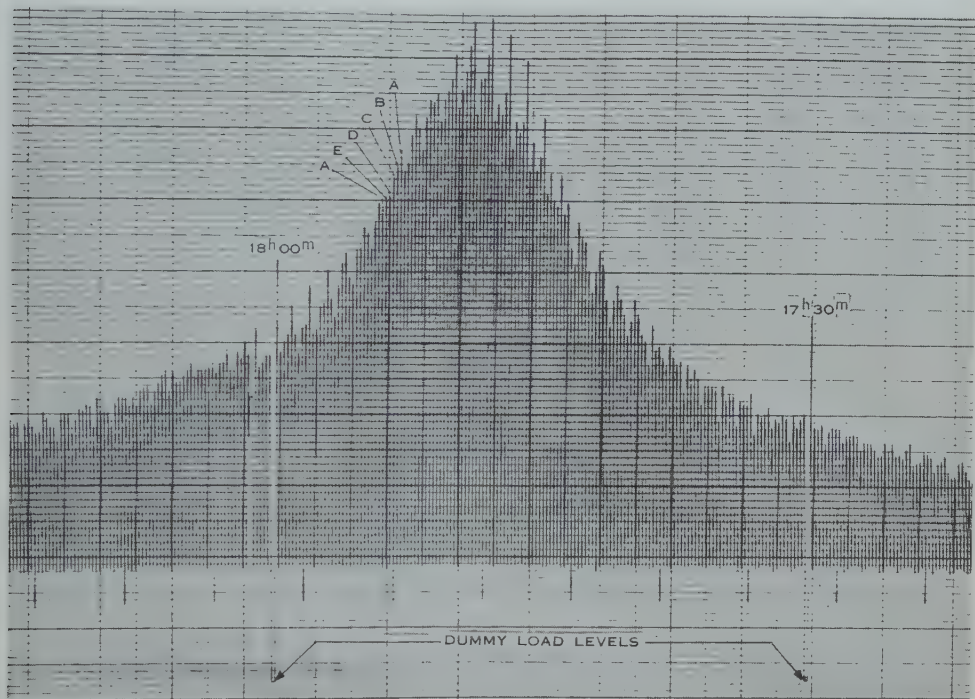


Fig. 1.—A record of the scanning type, described in paper I, used in the reductions of the present survey. This sample covers a period of nearly 1 hr when the beam crosses the galactic plane near $l=330^\circ$. The beam position is $N6$ ($\delta=-26^\circ 30'$) and the record was made on July 4, 1956. The two levels near the base of the record and immediately following the time marks are dummy-load levels. Several deflections indicating temperatures on different subpositions have been marked.

II. SCANNING OBSERVATIONS WITH THE 3.5 M "CROSS"

For the purpose of the survey, the aerial is used as a meridian transit instrument, so that material is collected from scans across the galactic plane. On each scan the aerial is switched electrically in a time interval of 1 min between five declinations spaced approximately 20 min of arc apart. The individual declination settings of each scan have been denoted by the letters *A*, *B*, *C*, *D*, and *E*, which we shall call subpositions. The declination of the *C* subposition defines the scan, or beam position, *A* and *E* being respectively the most southerly and northerly subpositions.

The range of declinations available to the aerial (from $\delta = +10^\circ 25'$ to $-78^\circ 7'$, corresponding to zenith distances from $+44^\circ 16'$ to $-44^\circ 16'$) is divided into 60 beam positions, 30 on each side of the zenith, labelled *N1* to *N30*, *S1* to *S30*. The zenith angle of each beam position is defined by equation (28) in paper I. These beam positions are arranged so that the *A* subposition of any beam position has the same declination as the *E* subposition of the next beam position to the south. This overlap in declination is useful for relating the calibrations of the individual beam positions, and, as noted later, allows the asymmetry in the aerial beam on these subpositions to be removed merely by averaging the two sets of temperatures observed on the adjacent beam positions at these declinations.

With the scanning system in operation, the recorder provides what is called a scanning record. A typical example of such a record during the transit of portion of the galactic plane through the aerial beam is shown in Figure 1. As described in paper I, the length of each sawtooth on the record represents the pencil-beam output at a particular declination and Right Ascension. The temperature of the north-south array, which must be added to the pencil-beam temperature to give brightness temperatures, is given by the deflection between the dummy load levels, two of which appear in Figure 1, and the base of the sawtooth pattern.

III. INSTRUMENTAL EFFECTS

Prior to describing the survey it will be advantageous to examine some features of the aerial and receiving systems in so far as they influence the interpretation and reduction of records. We consider firstly the aerial directivity effects.

(a) *The Aerial Directivity*

A qualitative plan-picture of the aerial pattern when the aerial is switched to a *C* subposition is shown in Figure 2 (a). It consists of a central pencil beam, the filled circle, around which is situated the side-lobe pattern. The largest side lobes are to be found distributed along the meridian plane which corresponds to the direction of the fan beam of the east-west array. Here the side lobes, except for several adjacent to the pencil beam, are randomly distributed and their responses average about 2 per cent. of the pencil-beam response. Another important group of side lobes lies along a small circle on the sky about an axis formed by the north-south arm of the cross. This small circle intersects the meridian plane at the declination of the pencil beam, and corresponds to the fan beam of the north-south array. The side-lobe structure here is similar to that on the meridian, but the average response is only about 0.5 per cent. of the pencil-beam response. These two groups of side lobes we call the primary side-lobe pattern, and they are represented by the shaded area in Figure 2 (a). The remainder of the diagram is covered by a random distribution of side lobes, whose responses are less than 0.02 per cent. of the pencil-beam response. These side lobes are too small to be of any consequence.

Since we are using a receiving system which is sensitive to the relative phases of the outputs from the two arms of the cross, side lobes will produce

recorder deflections which are of the same or opposite sign as that due to a radio source in the pencil beam. These are called positive and negative side lobes respectively. Randomness in the side-lobe structure arises because there are small departures from the nominal current amplitude and phase distribution along the arrays. However, the innermost primary side lobes are not of a random nature, but depend largely on the current distribution along the array.

As described in paper I, additional side lobes are introduced along the meridian plane when the beam is directed to the *A*, *B*, *D*, and *E* subpositions. These side lobes arise because displacement of the beam is effected by adjusting the phases of the north-south array dipoles in banks rather than individually; they are called switching side lobes. Near the pencil beam, these side lobes are spaced at intervals of $2^\circ.3$ for *B* and *D* subpositions, and several such side lobes

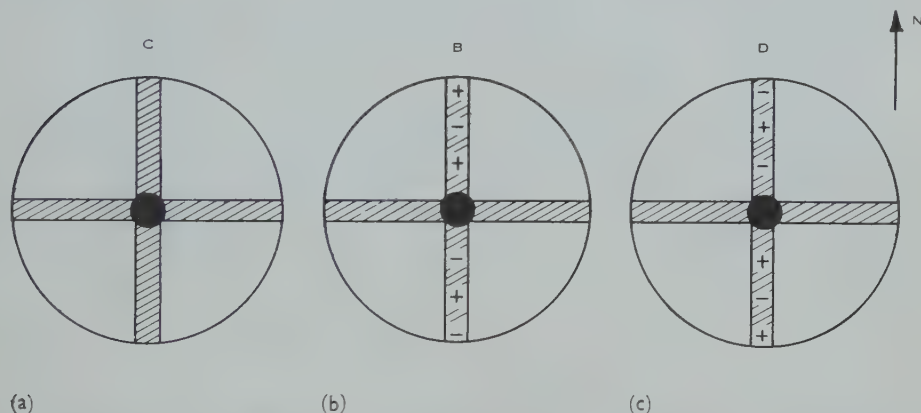


Fig. 2.—Pictorial representations of the aerial patterns of the "cross" on beam subpositions *C*, *B*, and *D* are shown in (a), (b), and (c) respectively. The central circle in each instance represents the pencil beam, whilst the shaded regions in the north-south and east-west directions represent the primary side-lobe pattern. The secondary side-lobe pattern fills the four unshaded sectors in each pattern. For the *B* and *D* positions, the positions and phases of switching side lobes are indicated qualitatively by the \pm signs.

in the neighbourhood of the pencil beam are shown in Figures 2 (b) and 2 (c) by $+$ and $-$ signs, which indicate the phases of the side lobes. Responses of switching side lobes diminish rapidly on either side of the pencil beam; the two innermost side lobes on each side of the pencil beam have relative responses 15 and 7 per cent. and provide the most important spurious contributions to brightness temperatures. Aerial patterns for *A* and *E* subpositions are similar to those of *B* and *D* subpositions, the chief difference being that near the pencil beam the switching side-lobe spacing is about $4^\circ.6$.

Switching side-lobe effects are present on all subpositions, except *C*, whenever the side lobes straddle an area of the sky in which there is a temperature gradient. They are therefore particularly noticeable in the region of the galactic plane. However, due to the fact that *A* and *E* subpositions overlap, and, further, that on these positions the switching side-lobe patterns are mirror images of one another, the average of *A* and *E* temperatures should contain no trace of their

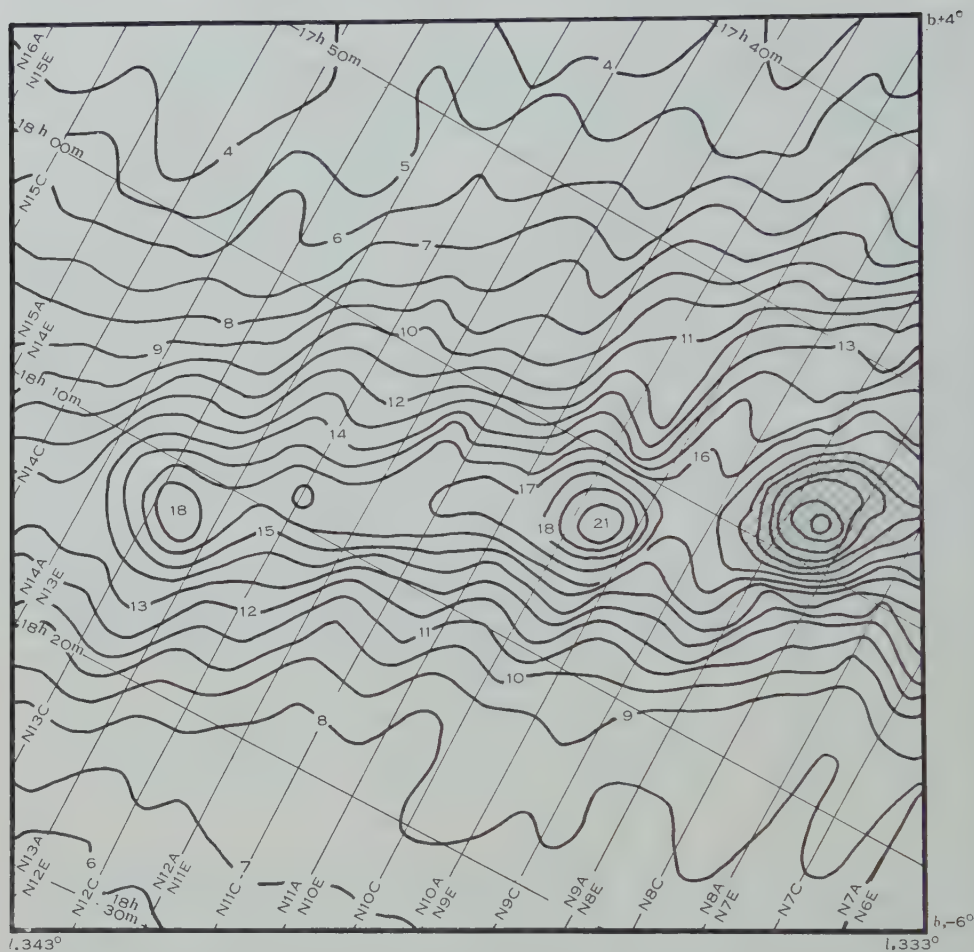


Fig. 3 (a).—Contours of brightness temperature for the strip $l=333$ to 343° before removal of beam asymmetry effects. The network of lines across the chart is the Right Ascension and declination grid, the latter being denoted by the beam subposition; lines marking declinations of B and D subpositions have been omitted. The contour interval is 1000°K except in the shaded region, where it is 2000°K .

effects. On the other hand, all B and D temperatures require individual corrections on this account. Switching side-lobe effects, which we call asymmetry effects, produce a waviness in the contours as shown in Figure 3 (a) for the $l=333^\circ$ to 343° , $b=+4^\circ$ to -6° region. This figure also illustrates the contour distortion effects of a strong source lying in switching side lobes. In the neighbourhood of R.A. $17^h 58^m$ on beam position $N12B$, the effect of source 17-2A (Mills 1952) in one of the 7 per cent. switching side lobes is seen clearly; stronger spurious responses of the same nature due to this source are seen on $N10B$ and D at the same Right Ascension, where it lies in a 15 per cent. switching side lobe. Contours of the same strip of the Galaxy, after correction for beam asymmetry, are shown in Figure 3 (b).

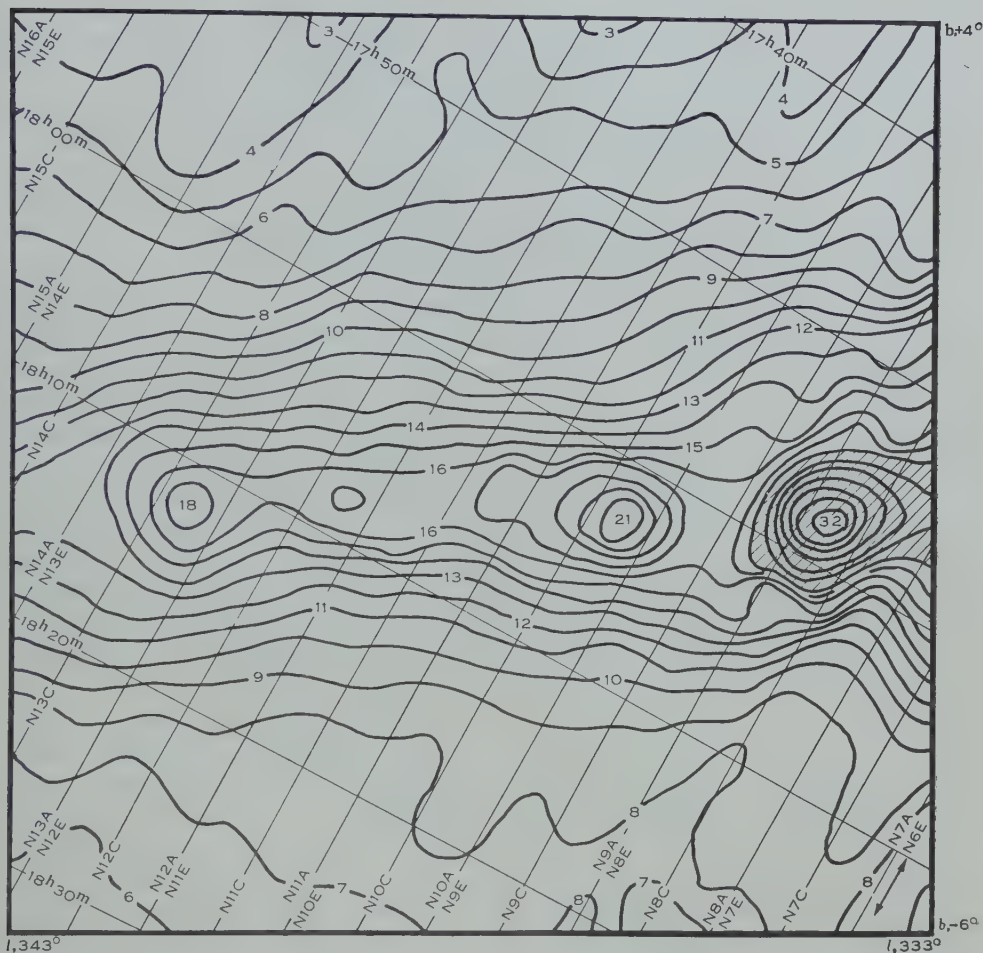


Fig. 3 (b).—Contours of brightness temperature for the strip $l=333$ to 343° after removal of beam asymmetry effects. Network and contour intervals as in Figure 3 (a).

For the determination of the switching side-lobe pattern, the reader is referred to paper I. There, it was shown that adequate corrections to B and D subposition temperatures are given by

$$\begin{aligned}\Delta T_{nB} &= 0.13(T_{(n-2)C} - T_{(n+2)E}) - 0.06(T_{(n-4)D} - T_{(n+3)B}), \\ \Delta T_{nD} &= 0.13(T_{(n+2)E} - T_{(n-2)C}) - 0.06(T_{(n+4)B} - T_{(n-3)D}),\end{aligned}$$

where suffixes nB and nD are the subpositions of beam position n at which temperatures are being corrected.

Finally, a strong source in one of the primary side lobes may affect brightness temperature measurements. The most serious of such effects is due to the intense central source of I.A.U. 13S4A (Centaurus-A) when it lies on the meridian. Adopting the average relative response of these side lobes, ± 2 per cent., this source could produce a modification of $\pm 1100^\circ\text{K}$ to the real brightness temper-

ature. Other bright sources in the field of the aerial could only produce effects up to ± 300 °K. Since the amplitudes and phases of the currents in the individual dipoles are not known sufficiently well to determine the side-lobe pattern, it is not possible to correct for these influences on measured temperatures. We will have cause to consider this effect in discussion of the contours (Section V).

(b) Calibration Changes

We have considered effects arising as the result of the reception of radiation in unwanted directions. The remaining instrumental effects are those related to variations in sensitivity or zero level as functions of time or beam position.

In order to obtain the brightness temperature of any region of the sky it is necessary, as described in paper I, to include a temperature derived from the north-south arm of the cross alone. But, as the aerial is switched between the five subpositions, small changes occur in the impedance presented by the north-south array to its receiver; their most serious effect is to cause changes in the zero level of the power output. Variations of this kind are responsible in part for the slightly irregular heights of the lower ends of the scans in Figure 1. When the north-south beam is directed close to the zenith the impedance changes become very marked because, as discussed in paper I, the standing-wave ratio along the feed line becomes high and needs special arrangements for cancellation; the degree of cancellation varies between different subpositions. For the two beam positions closest to the zenith (*S1* and *N1*), variations in zero level of up to 600 °K occur. Distortion in the contours arising in this manner have been neglected in the present survey. Since the brightness temperature always exceeds 5000 °K when the impedance changes are largest, the maximum possible error in the temperatures of two adjacent subpositions of the beam due to this cause is about 10 per cent., and over most of the area surveyed the error is very much less.

Absolute temperature measurements depend also on the stability of the receiver gain, which may be studied by means of the record calibration measurements. Although the calibration of the temperature scales was performed only at the beginning and end of each day's recording, lasting on the average 17 hr, a continuous check could be kept on the zero of the temperature scale by noting the variation of the dummy-load levels at half-hourly intervals. A study of the calibrations and dummy-load levels showed that the average change in sensitivity over this time interval was approximately 6 per cent., the change being essentially of a systematic nature. Hence we may conclude that the maximum receiver calibration error is unlikely to exceed 6 per cent. and, by making use of linear interpolations, this error was probably substantially reduced.

For the galactic plane studies a considerable amount of observational data was accumulated over the period from October 1955 to November 1956. In this interval two series of observations from beam positions *N30* to *S30* were carried out, yielding, after records free of obvious solar interference had been sorted out, one complete and another almost complete set of observations over the 60 beam positions. In each series the procedure was to observe firstly beam positions *N1* to *N30*, in that order, followed by adjustments to the arrays, if

necessary, after which observations were continued with southern beam positions *S1* to *S30*.

The lengthy period of observation has meant that consideration has to be given to the magnitude of long-term changes in the calibration. An estimate of such changes alone has not been attempted, but intercomparison of temperatures from duplicate records of the same beam position has been used to estimate the combined effects of short- and long-term changes. A detailed description of these investigations would be too tedious to give here, but the nature of the considerations involved will be indicated.

The comparison of the first and second series of observations has been restricted to the high temperature ridge lying near the galactic plane, principally because it is in this interesting region that an estimate of precision is most desirable. Therefore, for each record the average pencil-beam and north-south temperatures from the five consecutive beam subpositions giving the largest average brightness temperature were computed. The ratio of the average pencil-beam temperatures, and average north-south temperatures were then deduced from the two records generally available for each beam position.* Plots of these ratios are shown in Figures 4 (*a*) and 4 (*b*) as a function of beam position and the galactic longitude of the region where temperatures were measured. For some beam positions these figures indicate two ratios—in these instances the open circles identify the ratios with which we are concerned at present.†

Adopting the average of the two mean temperatures for each component as the basis for comparison of the two series of observations, it is found that the average deviation of the temperatures, both pencil-beam and north-south, is 5 per cent. The largest deviation is 15 per cent. This would indicate that averages of brightness temperature available for the area surveyed constitute a satisfactory combination of available observations. However, it was considered that a better compromise could be obtained near the galactic ridge by applying corrections to pencil-beam calibrations corresponding to points of Figure 4 (*a*) showing relatively large deviations. By so doing, it was hoped that the considerable fine structure apparent along and near the ridge-line would be defined more accurately. It should be noted that no weight was given to the north-south temperature ratios in deciding at what beam positions calibration adjustments were necessary, since the north-south contribution to the brightness temperature at the ridge is relatively small.

Closer scrutiny was therefore given to the original records showing the most serious discrepancies between the two measurements, namely, beam positions *N10* and *S1* to *S12*, and to their neighbouring beam positions. Since the pencil-beam temperatures for galactic crossings on *S13* to *S21* between galactic

* Since there were no records among the second series of observations covering the galactic crossings around $l=270$ to 295° on beam positions *S16* to *S20*, there are no temperature ratios in this range of longitudes.

† The fact that the more southerly declination circles cross the strip of the galactic plane with which we are concerned on two occasions explains the two sets of ratios possible for the more southerly beam positions.

longitudes 240 and 265° are for the most part too low to be corrected reliably, no attempt was made to modify their calibrations. By examining particularly the peak pencil-beam temperatures on the overlapping *A* and *E* subpositions in the vicinity of *S1* to *S12* it was found that much better continuity prevailed amongst the first series of observations on *N1A-S1E*, and *S12A* and *S13E* overlaps. Thus the systematic difference apparent on Figure 4 (*a*) appears to be due to some, as yet untraced, effect which caused the second series of temperatures on *S1* to *S12* to be about 20 per cent. too low. This disagreement has been removed by adding 20 per cent. to all the second survey pencil-beam and north-south temperatures on these beam positions—the latter temperatures being

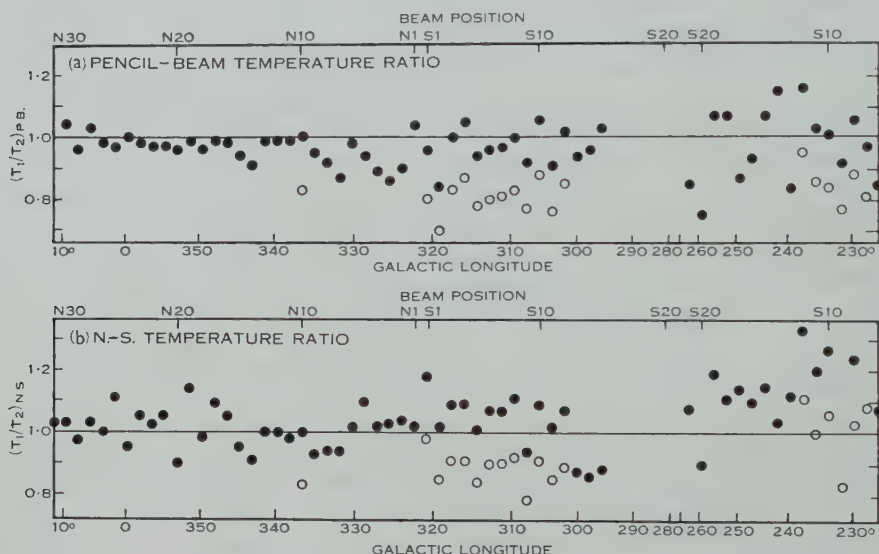


Fig. 4.—These diagrams depict the comparison of temperatures on the galactic ridge from the two observations available for nearly every beam position. (*a*) shows the ratio of the second to the first series of pencil-beam temperatures, (*b*) shows the similar ratio for the north-south temperatures. On beam positions *N10* and *S1* to *S12*, where calibration adjustments were made, the filled circles are based on the adjusted temperatures. Temperature ratios on these 13 beam positions, prior to calibration modifications, are shown by open circles.

included only because it was more convenient to correct total brightness temperatures than the pencil-beam contribution alone. The beam position *N10* was similarly treated—in this instance it was found that the first series of measurements was discordant. For all these 13 beam positions the ratio of second to first series observations, after the corrections just considered, have been indicated by the dots. It will be noted that in applying the corrections to north-south temperatures, the agreement between the two measurements for *N10* and *S1* to *S12* (from $l=300$ to 320°) has been improved. On the *S1* to *S12* ($l=227$ to 238°) crossings the agreement between the two series of north-south temperatures has deteriorated. This is not considered serious, since the average north-south temperature is only about 1500 °K.

To summarize, it has been found that for all beam positions, other than *N10* and *S1* to *S12*, an average of the two sets of observations is quite adequate. On the remaining 13 beam positions, for which agreement between our two series of measurements was not quite so good, appropriate corrections have been made to measurements of one series or the other to produce more suitable averages of the observations.

As a result of these corrections the average uncertainty in the relative brightness temperatures combined in the above manner should be less than 5 per cent. However, in respect to the absolute accuracy of brightness temperatures, the accuracy of the aerial gain measurements must be considered. Little (1958) considers that the uncertainty in the aerial gain is 10 per cent.; this means that the absolute accuracy of the survey temperatures would be somewhat better than 15 per cent.

Finally, it might be mentioned that, although the investigations outlined above have been restricted to the neighbourhood of the galactic ridge, the conclusions regarding the accuracy of brightness temperatures will apply all over the area of sky considered. The principal reason for this is that for the most part the aerial scans across the galactic plane region between $b = +4$ and -6° in so short a time that there is little chance of significant variation from the values taken at the crossing itself.

IV. REDUCTION OF DATA FOR THE SURVEY

As was mentioned earlier, the region of the galactic plane to be considered comprises the strip between latitudes $+4$ and -6° and stretches from $l = 223^\circ$ through the galactic centre to $l = 13^\circ$.*

In collecting material for this area, brightness temperature measurements were made from tracings of the original records prepared in a manner described in paper I. As in the preliminary analysis described in Section III (*b*), only records not obviously affected by solar interference have been used. Temperature scales for the north-south and pencil-beam components were obtained by linear interpolation from the calibrations at the beginning and end of each record. Temperatures were measured to within $\pm 150^\circ\text{K}$ at intervals of 2 min in Right Ascension on every subposition of the beam. In complex regions of the sky, temperatures were measured at 1 min intervals. After the corrections indicated by the preliminary comparison (see Section III (*b*)) were applied to these measurements, the results of the two series of observations were averaged. In regard to this averaging it might be added that not only are instrumental effects reduced, but any less obvious effects of solar interference on the records is also diminished. Further, averaging of *A* and *E* subposition temperatures removed beam asymmetry effects at these declinations.

* No data are obtainable for longitudes from 13 to 167° , since these regions lie to the north of the normal operating declination range of the aerial. Here the aerial efficiency is a rapidly decreasing function of zenith angle and the relative importance of spurious responses increases. The strip from $l = 167$ to 223° has been subjected to a preliminary investigation which showed that this comparatively low temperature region would require special methods of reduction if features present were to be described accurately. We have consequently deferred consideration of this region until a later date.

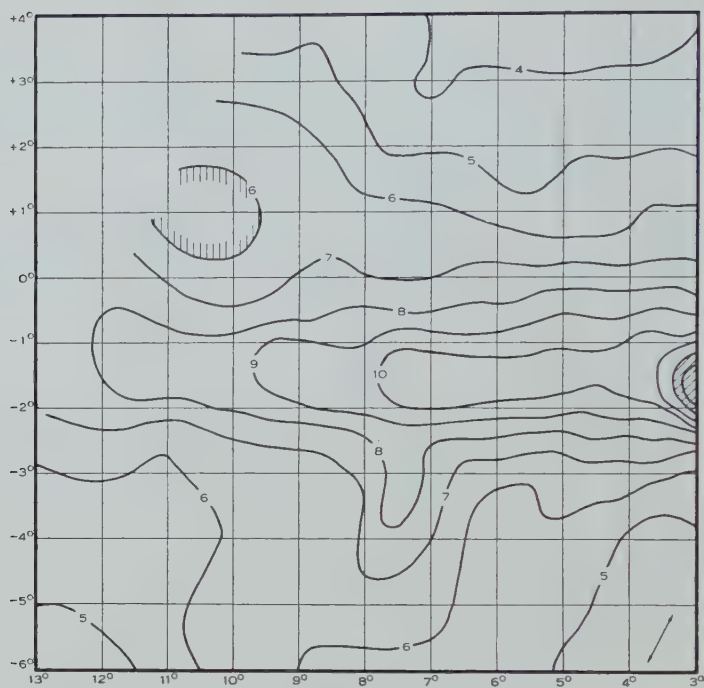


Fig. 5 (a).—For explanation, see Section V.

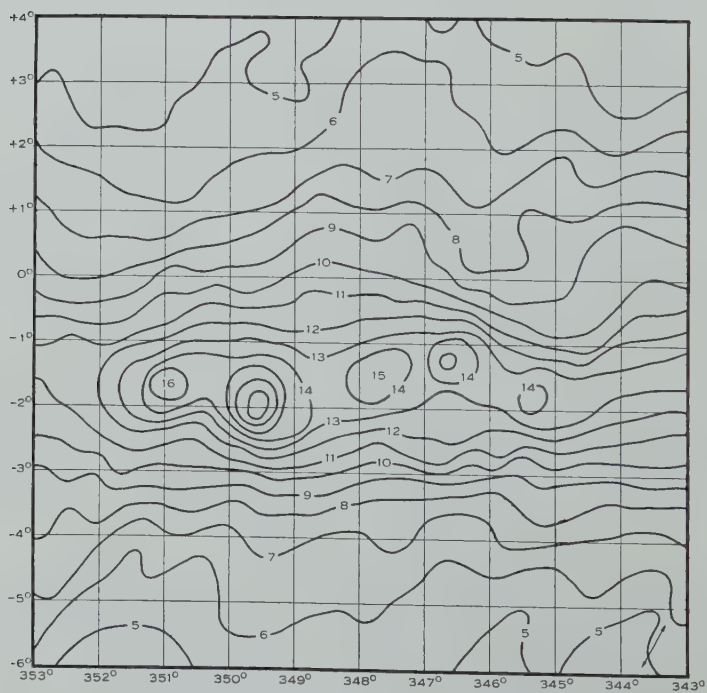


Fig. 5 (c).—For explanation, see Section V.

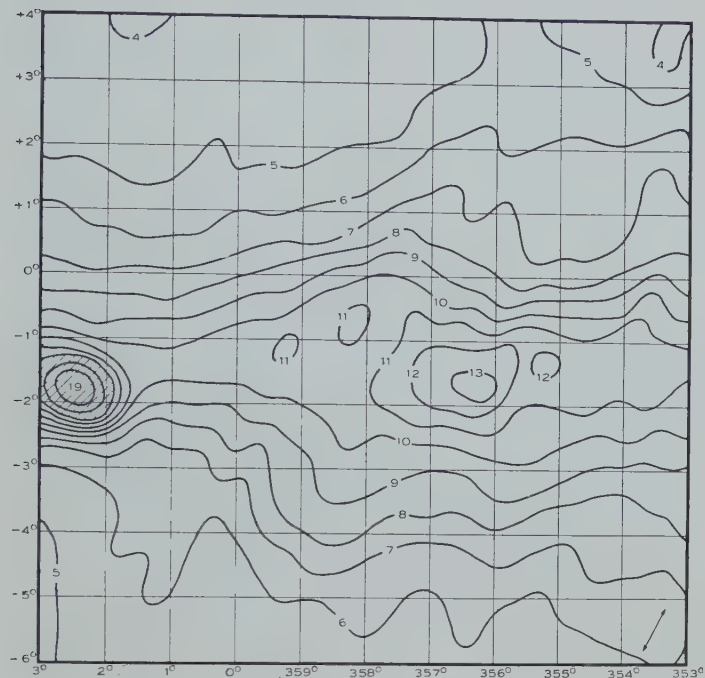


Fig. 5 (b).—For explanation, see Section V.

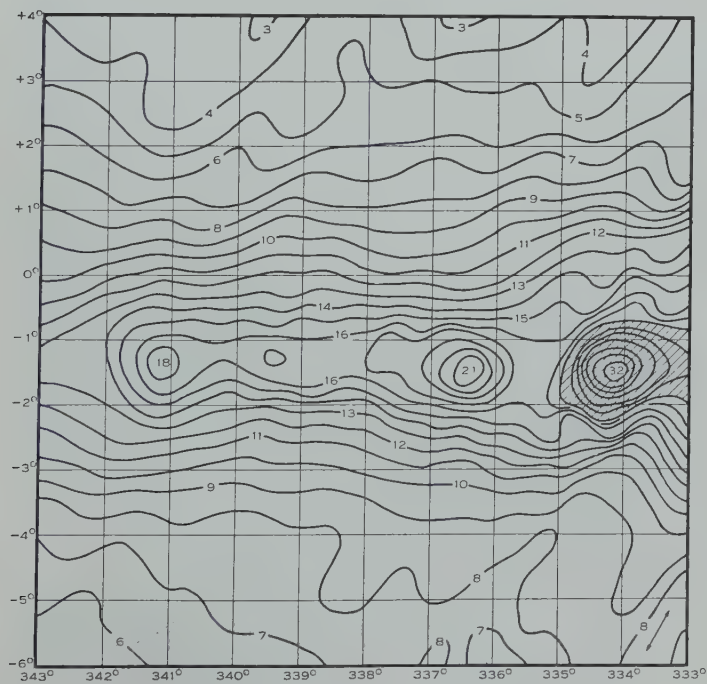


Fig. 5 (d).—For explanation, see Section V.

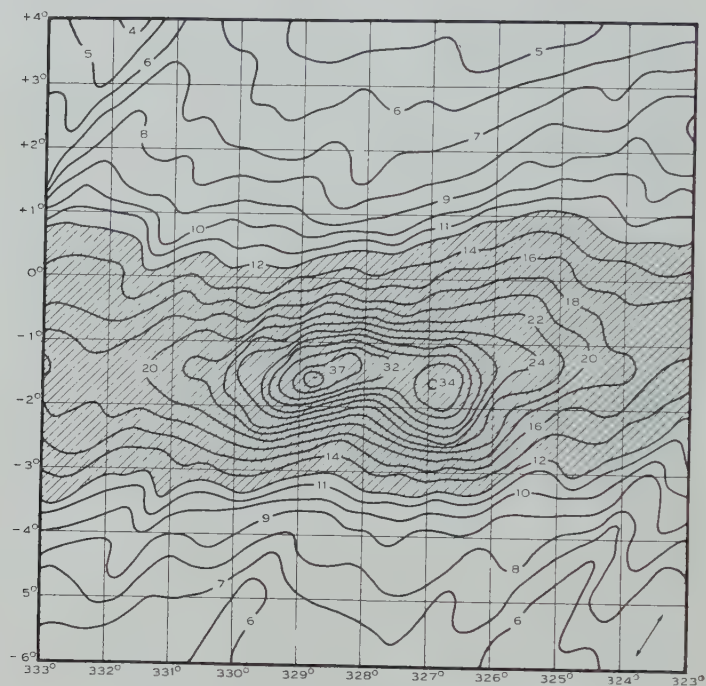


Fig. 5 (e).—For explanation, see Section V.

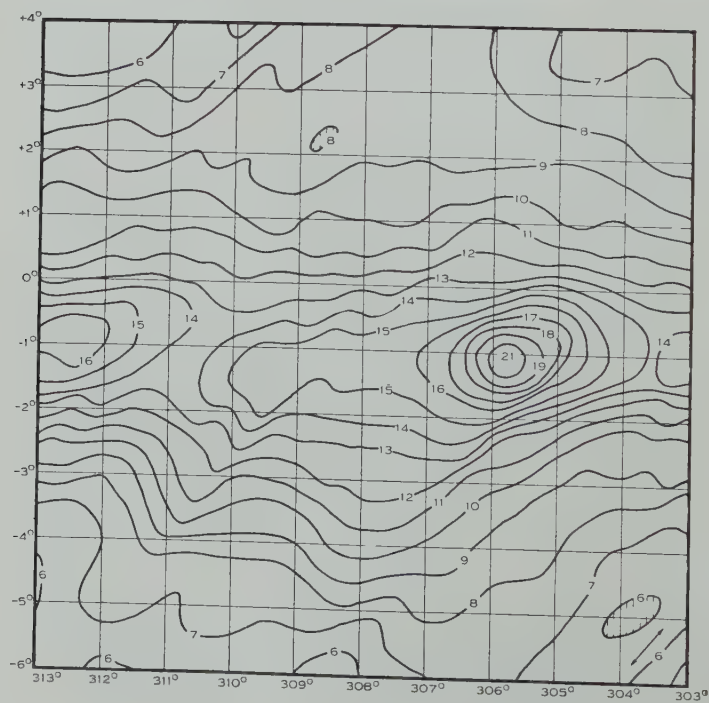


Fig. 5 (g).—For explanation, see Section V.

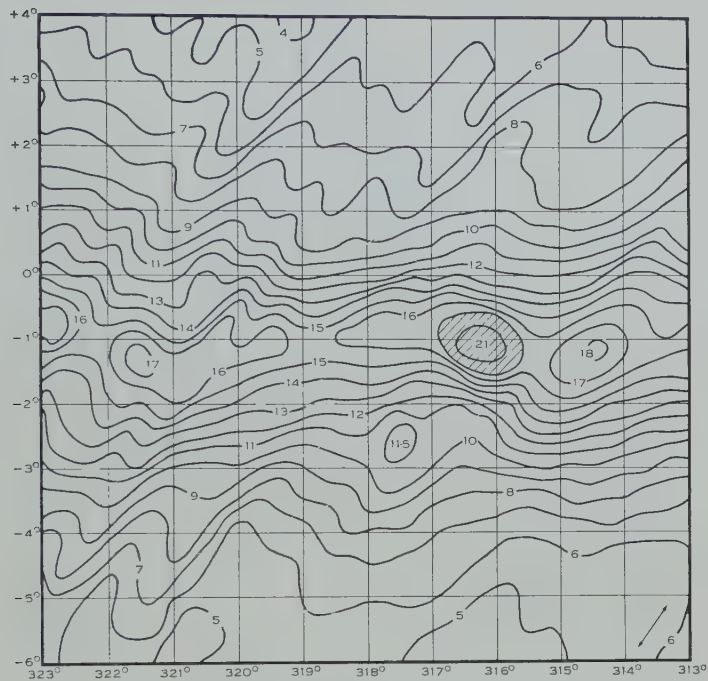


Fig. 5 (f).—For explanation, see Section V.

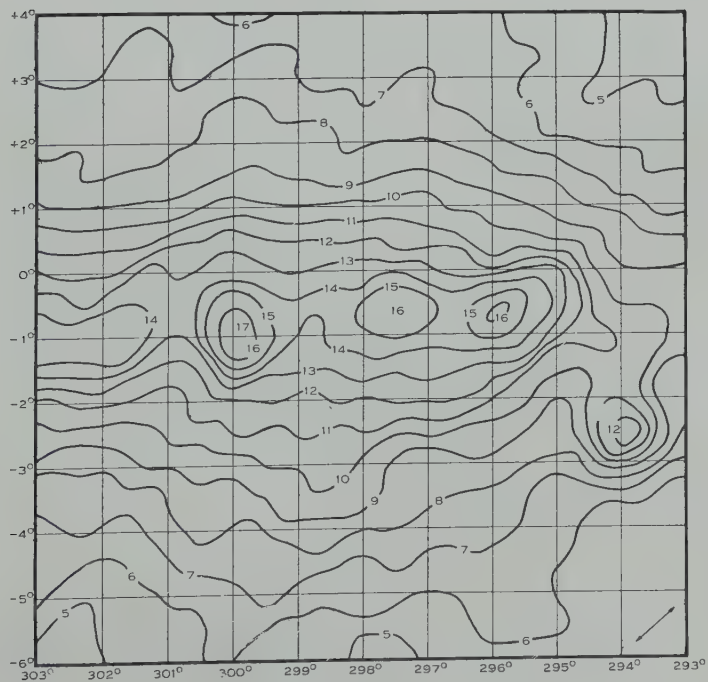


Fig. 5 (h).—For explanation, see Section V.

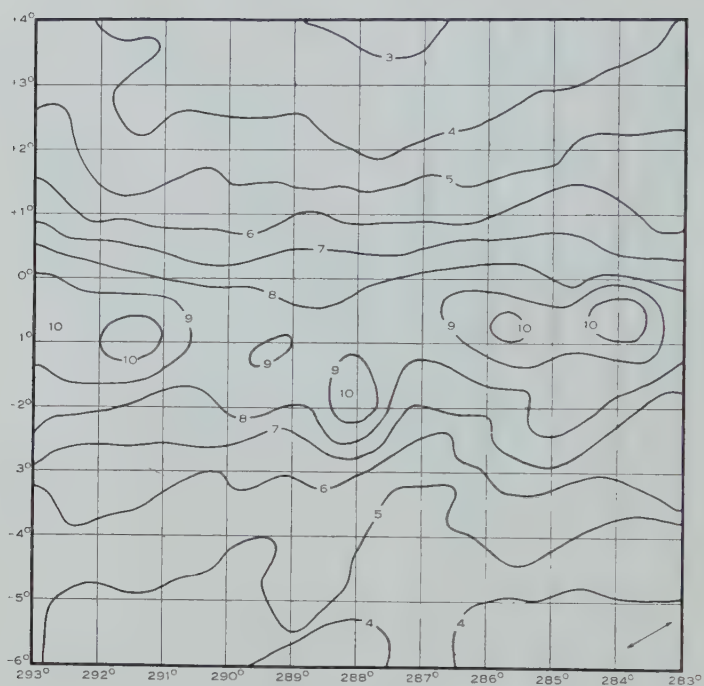


Fig. 5 (i).—For explanation, see Section V.

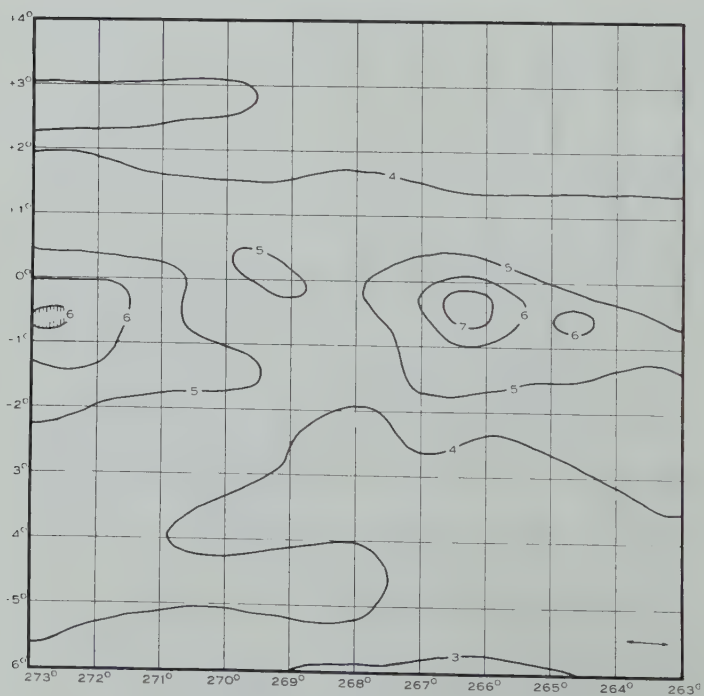


Fig. 5 (k).—For explanation, see Section V.

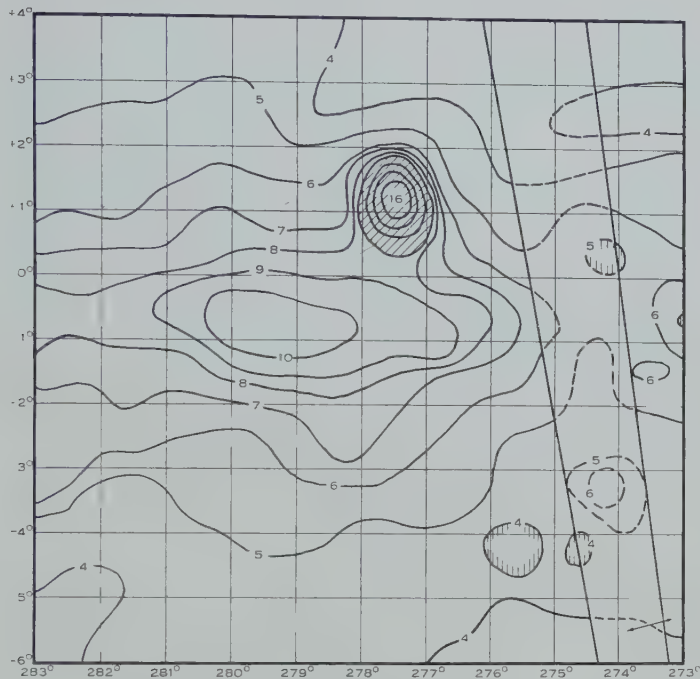


Fig. 5 (j).—For explanation, see Section V.

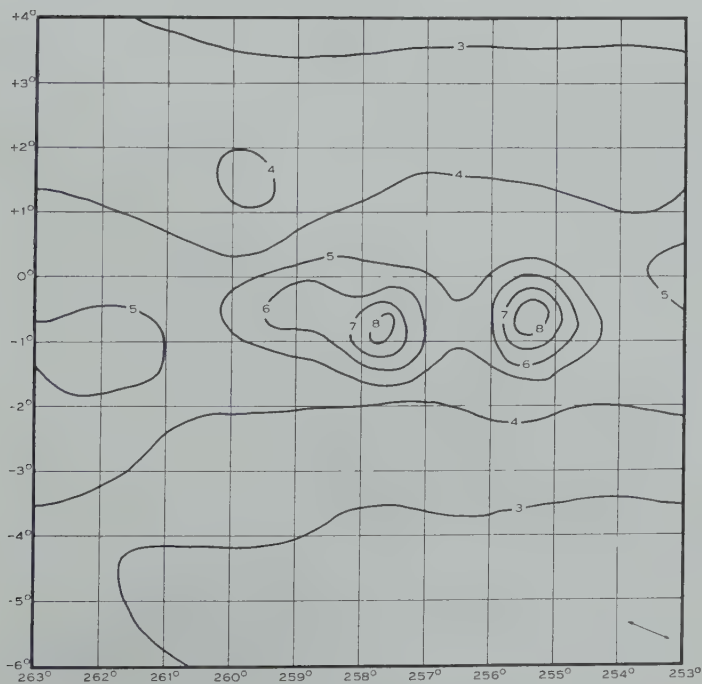


Fig. 5 (l).—For explanation, see Section V.

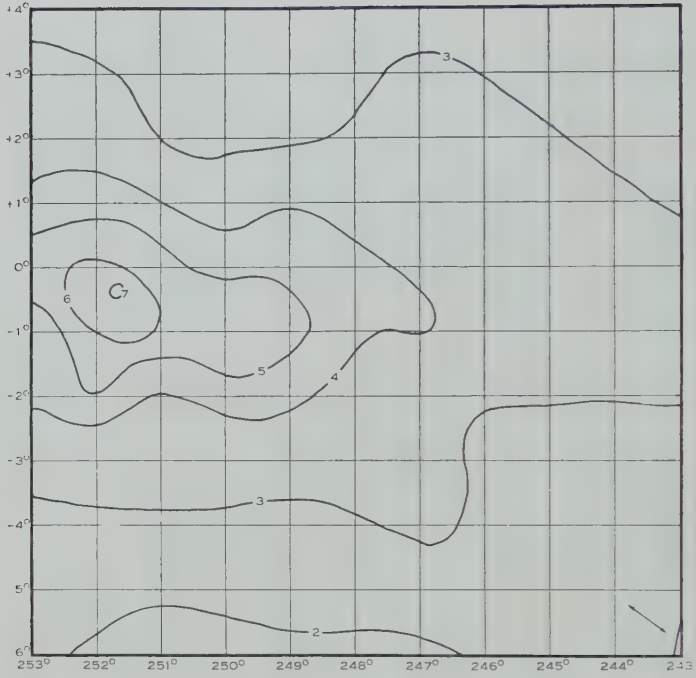


Fig. 5 (m).—For explanation, see Section V.

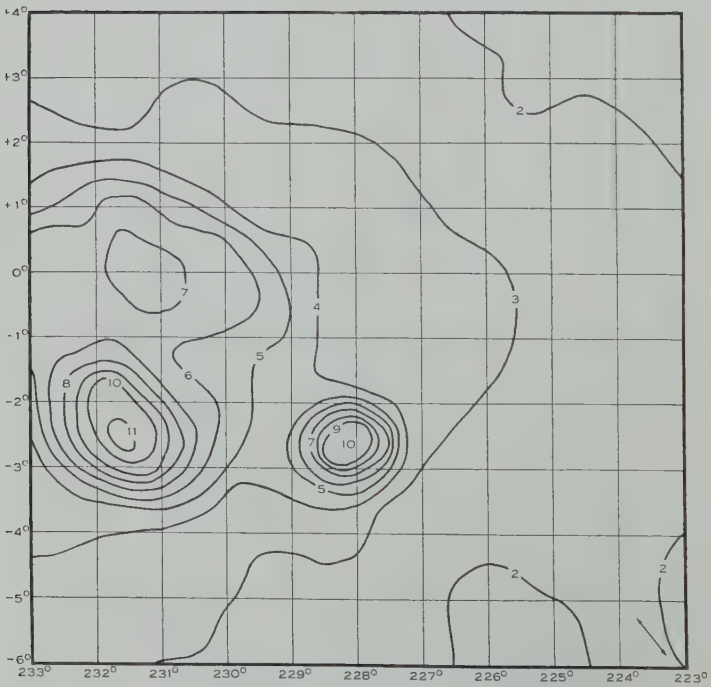


Fig. 5 (o).—For explanation, see Section V.



Fig. 5 (n).—For explanation, see Section V.

Charts, each 10 by 10° in galactic latitude and longitude, were then prepared with a Right Ascension and declination grid such that each intersection of the grid corresponded to a direction in which temperatures had been measured from the tracings. The construction of the grid was based on data prepared in the Laboratory by Martha Stahr Carpenter from the Lund tables with precession corrections necessary to render the conversion valid for epoch 1955. This being only about a year and a half from the mean epoch of observation meant that, since the grid was intended to be accurate to 0.1° in l and b , adjustment of the grid to make it appropriate for the mean epoch of observation was not warranted. Average temperatures as deduced above were then inserted at their appropriate Right Ascension-declination grid junction.

The beam asymmetry corrections for all B and D subpositions were then computed from the relations noted earlier (Section III (a)) using the average temperatures inserted on the charts. Although this involved in many cases using uncorrected B and D temperatures, the effects produced in doing so are of minor importance and were neglected. Beam asymmetry corrections around $b = +4$ and -6° were always less than several hundred degrees, irrespective of galactic longitude. Around the galactic plane, however, larger corrections occur because of the greater background temperature gradients. This is especially so about the galactic centre where $\pm 1000^\circ\text{K}$ corrections were common, whilst in small areas in this locality corrections of up to $\pm 2000^\circ\text{K}$ have been necessary. Generally, however, the asymmetry corrections were less than $\pm 1000^\circ\text{K}$.

From the final temperature grid on the charts, contours were constructed—these are shown in Figures 5 (a) to 5 (o).^{*} No use has been made of Bracewell's (1956) process of filtering out artificial high frequency components since the method is too laborious for manual reductions and, moreover, the uncertainties in the data do not warrant such a detailed analysis. Finally, no correction has been made to the contours for the finite width of the pencil beam.

V. DISCUSSION OF CONTOUR DIAGRAMS

Contours of brightness temperature in Figures 5 (a) to 5 (o) constitute the results of the survey. Temperatures marked on these diagrams are in units of 1000 °K, which is also the interval of the contours except in the diagonally shaded regions where the interval is 2000 °K. Contours enclosing cool regions have vertical shading on the low temperature side of the contours. Where the contour system is inadequate for the purpose of illustrating peak temperatures of sources, or in extended bright areas, extra temperatures have been added. These temperatures are not embraced by contour lines. In the lower right-hand square of the l - b grid on each chart the average direction of the aerial scan is indicated by the double-headed arrow.

A general inspection of the diagrams indicates that the beam asymmetry corrections have successfully removed most of the spurious waves in the contours. There are traces of asymmetry effects remaining, for example, around the galactic centre region $l=323$ to 333° , but they do not influence the interpretation sufficiently to warrant more precise correction.

There are several spurious effects remaining in the contours to which attention should be drawn. Firstly, there is the region wherein the effects of Centaurus-A in the north-south part of the primary side-lobe pattern may produce significant distortion of the contours. The narrow strip which is affected occurs around $l=274$ – 275° . This strip has been indicated on Figure 5 (j) and contours within it have been broken. In this uncertain region, there is at $l=274\frac{1}{4}^\circ$, $b=-3\frac{1}{3}^\circ$ a temperature of about 2000 °K above the surroundings. Whether this is a real source, or whether it is one of the rare examples of a relatively large primary side lobe cannot be ascertained. The side lobe would be required to have a response of 4 per cent. of the primary beam to produce this effect as compared with an estimated r.m.s. primary side-lobe response in the north-south direction of 1.7 per cent.

Another area where spurious effects occur lies close to the zenith in a strip lying diagonally across the charts between $l=323^\circ.1$ and $326^\circ.2$ at $b=-6^\circ$, and $l=316^\circ.4$ and $319^\circ.7$ at $b=+4^\circ$. Here the impedance effects, mentioned earlier, have their maximum influence and introduce spurious detail into the contours.

A map of contours obtained with this instrument in the region of the galactic centre, the area corresponding to our Figure 5 (e), has been published by Mills (1956). Although many of the records used in these two maps were the same,

^{*} A large diagram, measuring 6 by 90 in., in which the contents of Figure 5 are combined into one continuous chart, has been prepared, and a limited number of copies are available on request.

there are some differences in the published results, the most notable being the higher brightness temperatures obtained now. These result from a recalibration of the equipment as described in paper I. Other small and relatively unimportant differences in the shapes of the contours arise from the use of different reduction processes and the use of additional observational material in the present map. While this map is, in general, more reliable, the source at the galactic centre was treated more carefully in the earlier investigation, using extra records at intermediate beam positions: these were not used in the present reduction as the aim was to produce a uniform and more or less standardized treatment of the whole area covered by the maps.

Another and much larger area has been treated more carefully in an independent investigation by H. Rishbeth (1958); this includes the bright regions around longitude 230° as well as adjacent areas at higher latitudes not included in our maps. It is probable that other localized areas will also receive special treatment in subsequent analyses.

VI. ACKNOWLEDGMENT

The authors wish to express their indebtedness to Mrs. M. Johnson for the conscientious manner in which she carried out her share of the laborious reductions involved in this work.

VII. REFERENCES

- BOLTON, J. G., WESTFOLD, K. C., STANLEY, G. J., and SLEE, O. B. (1954).—*Aust. J. Phys.* **7**: 1.
BRACEWELL, R. N. (1956).—*Aust. J. Phys.* **9**: 297.
LITTLE, A. G. (1958).—*Aust. J. Phys.* **11**: 70.
MILLS, B. Y. (1952).—*Aust. J. Sci. Res. A* **5**: 266.
MILLS, B. Y. (1956).—*Observatory* **76**: 65.
MILLS, B. Y., HILL, E. R., and SLEE, O. B. (1958).—*Observatory* **78**: 116.
MILLS, B. Y., and LITTLE, A. G. (1953).—*Aust. J. Phys.* **6**: 272.
MILLS, B. Y., LITTLE, A. G., SHERIDAN, K. V., and SLEE, O. B. (1958).—*Proc. Inst. Radio Engrs., N.Y.* **46**: 67.
RISHBETH, H. (1958).—*Aust. J. Phys.* **11**: 550.
SCHEUER, P. A. G., and RYLE, M. (1953).—*Mon. Not. R. Astr. Soc.* **113**: 1.

RADIO EMISSION FROM THE VELA-PUPPIS REGION

By H. RISHBETH*

[Manuscript received July 28, 1958]

Summary

The Vela-Puppis region has been surveyed with Mills Cross radio telescopes, and radio isophotes plotted at two wavelengths. Many discrete sources have been observed. Several H II regions have been detected in emission at 3.5 m and in absorption at 15.2 m; the radiation from such objects is of thermal origin. The outstanding feature of this region is a group of strong non-thermal sources near the galactic equator. These are superimposed on an intense belt of radiation along the equator, and this, too, is certainly non-thermal in origin. Some correlation with galactic structure is suggested.

I. INTRODUCTION

The constellations of Vela and Puppis contain some of the brightest emission regions in the southern Milky Way. The most extensive H II region known lies in this area; it is excited by the early-type stars γ Velorum and ζ Puppis, and its fragments are scattered over several hundred square degrees of sky (Gum 1955, 1956). The present paper describes general surveys of the Vela-Puppis region at two radio wavelengths. These surveys have shown that some parts of this H II region can be detected in emission or absorption, depending on the wavelength used, but they have also revealed a compact group of non-thermal sources near the galactic equator, and the characteristics of these sources are studied in some detail.

The instruments used were the two Mills Cross radio telescopes at the Fleurs Field Station, near Sydney, N.S.W. (151° E., 34° S.). In each system the outputs of two long mutually perpendicular arrays are combined in a phase-switching receiver to give a narrow primary beam which can be pointed to any direction in the meridian by altering the phasing of the dipoles in the north-south array. In normal use the beam is not swung more than about 45° from the zenith. The aerial beams are circular and of Gaussian cross section, with half-power widths 0°·8 and 1°·4 at 3.5 and 15.2 m respectively for observations not far from the zenith. The aerial systems and their associated equipment are described in separate papers (Mills *et al.* 1958; Shain 1958).

II. OBSERVATIONS AND RESULTS

(a) Observations at 3.5 m

A complete set of records covering the sky bounded by R.A. 0600 and 1000, Dec. -34° and -57°, was used. These records are of the "scanning" type previously described (Mills and Slee 1957) in which five neighbouring

* Division of Radiophysics, C.S.I.R.O., University Grounds, Chippendale, N.S.W.; present address: Cavendish Laboratory, Cambridge.

declinations are surveyed quasi-simultaneously; the north-south aerial of the cross is switched to each declination for 12 sec in every minute. The interval between these positions is about 20', and the subsequent record is taken with the central declination displaced about $1^{\circ} 20'$ from its predecessor; the most northerly scan of one record (known as *E*) thus coincides with the most southerly or *A*-scan of the next.

For the region of the sky given above, the values of equivalent aerial temperature T_a were plotted on a grid in galactic coordinates, points being taken on each scan at intervals of 2^m in Right Ascension (1^m in regions of special interest). The *A*-scan of each record was combined with the *E*-scan of the next, and, where a strong north-south gradient of aerial temperature occurs, as in low galactic latitudes, small corrections for beam asymmetry were applied to the *B* and *D* positions. The values of T_a were plotted to the nearest 50°K —this representing the fluctuation level of the records—and contours drawn at intervals of 250 or 500°K . The resulting map is shown in Figure 1; closed contours around sources of small angular extent have been omitted.

(b) *Observations at 15.2 m*

The operation of the long-wavelength cross is curtailed by the ionosphere by day and by interference from broadcasting stations at night. During November and December 1956, when the majority of the observations were taken, records of reasonable quality could only be obtained between about 0200 and 0600 hr, local time, using a bandwidth of 10 kc/s. The use of the aerial for other purposes precluded observation before about 0715 sidereal time, so the preceding parts of the principal Vela-Puppis H II region were missed; otherwise a reasonable coverage of the area between Dec. -39° and -52° was achieved, with the aid of a few records taken in November 1957 to fill in some of the missing portions. Values of T_a were plotted at intervals of 3^m in Right Ascension, and a chart of the region in galactic coordinates derived (Fig. 2), with contours at intervals of $30,000^{\circ}\text{K}$.

Observations at this long wavelength are subject to errors due to ionospheric refraction. Besides the expected refraction for a spherical ionosphere, routine observations of intense discrete sources show northward displacements of the order of 1° , due to the prismatic effect of the north-south gradient of total *F*-region ionization (Shain, personal communication). As all the records were taken at about the same season and time of day, a uniform correction in declination of $\frac{1}{2}^{\circ}$ has been applied to the whole chart; this is consistent with the observed position of the source Puppis-A.

(c) *Discrete Sources*

A search for discrete sources was made on each of the two grid plots of equivalent aerial temperature. When the possible features appearing on at least two adjacent scans had been listed, they were checked on the tracings of records from which the grid plots had been derived, and about half on each list rejected. Special attention was paid to positions in which side-lobe effects of intense sources might be expected to occur, and objects in such locations were

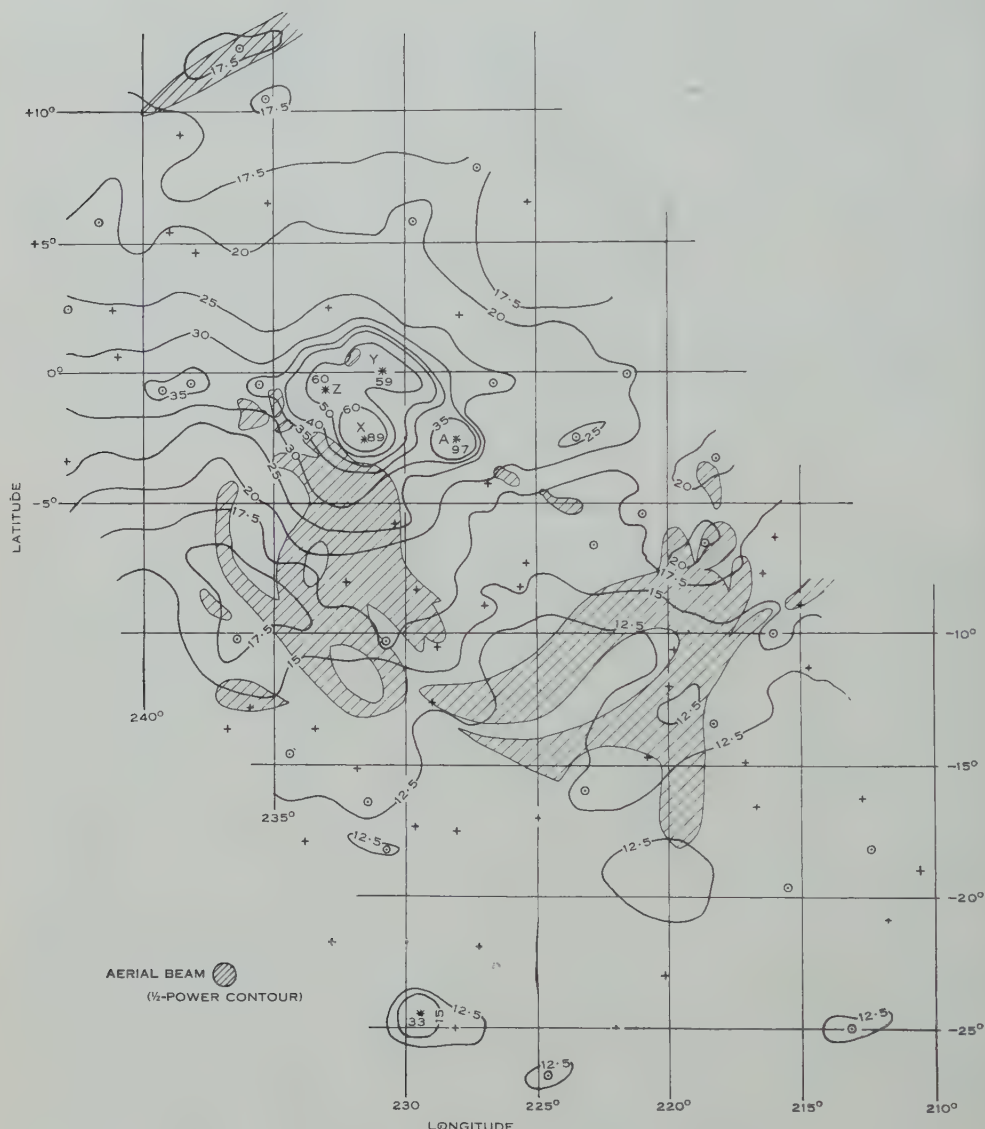


Fig. 1.—Contours of equivalent aerial temperature at 3.5 m wavelength. The unit is 100 °K, and the half-power contour of the aerial is shown in the bottom left-hand corner. The shaded areas show the extent of H II emission from the large, nearby H II region Stromlo 12 (after Gum 1956). The positions of discrete sources are also indicated on the figure. Sources of small angular size are indicated by crosses (+) and those with angular sizes comparable to the aerial beamwidth by circled points (⊙). Very closely spaced contours around the most intense sources have been omitted, the peak temperature only being given. The sources Puppis-A, Vela-X, Vela-Y, and Vela-Z are distinguished by the letters A, X, Y, and Z respectively.

retained only after severe scrutiny of the tracings. Finally, the positions and intensities of the remaining sources included in the list were measured from the tracings.

In Appendix I, Table 1 lists 77 sources observed at 3.5 m.* Each source is assigned a serial number with prefix R, and a reference indicating galactic position, to the nearest degree, is also shown. The third and fourth columns give the

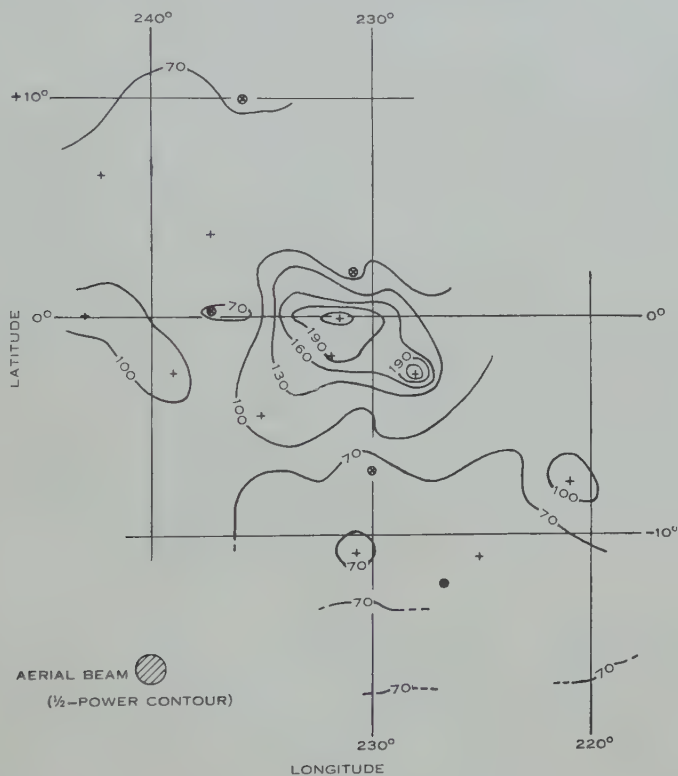


Fig. 2.—Contours of equivalent aerial temperature at 15.2 m wavelength. The unit is 1000 °K, and the half-power contour of the aerial beam is shown. Discrete sources (+) and localized minima due to absorption (⊗) are indicated.

Right Ascension and declination at epoch 1950. The accuracy of the estimated positions varies from source to source. For the strongest, clearly defined sources the probable errors are estimated to be about $0^m.3$ in Right Ascension and $3'$ in declination. The uncertainty is increased for the weaker sources ($0^m.5$ and $10'$), whilst the probable errors may be up to $1^m.0$ and $20'$ for very extended, or possibly confused, objects. The fifth column shows ΔT_a , the increase in equivalent temperature due to the source, and for sources which do not appear

* A definitive catalogue of discrete sources observed with the 3.5 m Cross is being prepared by Mills and his collaborators, and this catalogue will, in due course, include the Vela-Puppis region.

to be appreciably extended the corresponding flux density S is given in units of $10^{-26} \text{ W m}^{-2} (\text{c/s})^{-1}$. For a few bright extended sources, an integrated flux density has been estimated and inserted in column 6 in italics. In the final column, a bracketed reference to a strong source, e.g. (Hya-A), indicates a possibility of interference from that source; although, as mentioned before, most of the objects of which the reality is doubtful have been omitted from the list. Also in this column are references to previous source surveys and possible identifications with optical objects, the more doubtful identifications being indicated again by brackets.

Observations at 15.2 m show localized minima as well as maxima, and Table 2 gives details of the 16 discrete features observed at this wavelength. These are allotted three-figure serial numbers, with the localized minima, or negative sources, numbered in italics. These are included only if the absorption is sufficiently great to produce actual minima in T_a ; in fact absorption might be expected to occur over considerable areas of the sky. Owing to ionospheric refraction and the poorer signal-to-noise ratio at this wavelength, the quoted positions of the sources in Table 2 are less accurate than those in Table 1. The probable error in Right Ascension is at least 1^{m} for the brightest sources and 2^{m} for the others. The declinations of all objects have been shifted, as described in Section II (b), and the probable errors in this coordinate are from $\frac{1}{3}^\circ$ for the brightest up to 1° for the fainter or extended objects. ΔT_a is given in units of 1000°K , and the corresponding peak flux densities in units of $10^{-26} \text{ W m}^{-2} (\text{c/s})^{-1}$; no attempt has been made to derive integrated flux densities. Possible correlations with Table 1 and with optical objects are indicated.

III. THEORY OF H II REGION OBSERVATIONS

According to standard theory (Piddington 1951; Mills, Little, and Sheridan 1956) the optical depth τ of a cloud of ionized gas radiating thermally is proportional to the integral along the line of sight of the square of the electron density N_e ; this integral is known as the emission measure (E) and its units are pc-cm^{-6} . For an electron temperature T_e of 10^4°K , the relationships between E and τ are given by

$$\left. \begin{aligned} E &= 1.6 \times 10^4 \tau & (3.5 \text{ m}), \\ E &= 8 \times 10^2 \tau & (15 \text{ m}). \end{aligned} \right\} \dots\dots\dots (1)$$

If an H II region is in front of a uniform background of brightness temperature T_b , its apparent excess temperature is

$$\Delta T = (T_e - T_b)(1 - e^{-\tau}), \dots\dots\dots (2)$$

and it appears in emission if $T_e > T_b$ and in absorption if $T_e < T_b$. For an optically thin region $\tau \ll 1$ and

$$\Delta T \simeq (T_e - T_b)\tau.$$

This approximation holds at 3.5 m for the H II regions to be considered in the Vela-Puppis region. Here, T_b is of the order of $2 \times 10^3^\circ \text{K}$, and thus, for $T_e = 10^4^\circ \text{K}$, the object is seen in emission with

$$E = 2.0 \Delta T \quad (3.5 \text{ m}). \dots\dots\dots (3)$$

This relation is not sensitive to changes in T_e or T_b . The constant on the right-hand side is proportional to $T_e^{3/2} \cdot (T_e - T_b)^{-1}$, so that, for the conditions just stated an increase of 20 per cent. in T_e changes the constant in equation (3) by only 5 per cent.

At 15 m T_b is of the order of 10^5 °K and the nebula is seen in absorption ($\Delta T < 0$). If the region is optically thin the relation corresponding to (3) above is

$$E \simeq 10^{-2} \Delta T \quad (15 \text{ m}), \quad \dots\dots\dots (4)$$

but the opacity of all but the faintest H II regions detectable optically ($EM \simeq 300$) is too great for this formula to be valid. In this case equation (2) must be used without approximation.

In general, the observed change in equivalent temperature, ΔT_a , will not be equal to ΔT , since there is some emission from the space between the H II region and the observer. The value of T_b refers only to the part of the received noise which has been generated behind the H II region. In the case of the H II regions considered here and for the 3.5 m observations, the small corrections required in the calculation of emission measures are negligible. On the other hand, for the 15 m observations the effect of emission from in front of the H II region cannot be ignored, as will be clear from a particular example discussed in Section IV.

Furthermore, the theory given above refers only to H II regions whose angular sizes exceed the aerial beamwidth. In other cases $\Delta T_a < \Delta T$ and the ratio of the quantities can usually be estimated only roughly. For objects of small angular size an integrated emission can be derived from the radio flux density, using the relations

$$S = 2k\lambda^{-2} \int \Delta T d\Omega \propto \int E d\Omega, \quad \dots\dots\dots (5)$$

provided $\tau \ll 1$. λ and k are the radio wavelength and Boltzmann's constant respectively and $d\Omega$ is an element of solid angle.

IV. DISCUSSION

(a) H II Regions

A sketch of the principal H II regions in Vela-Puppis has been given by Gum (1955, 1956) and the relevant part has been reproduced in Figure 3. Apart from a few compact nebulosities near the galactic plane, all the emission objects belong to the vast complex excited by the stars γ Velorum and ζ Puppis, assigned the number 12 in the Mount Stromlo catalogue (Gum 1955) and at a distance from the Sun of about 250 pc. The parts of Stromlo 12 in the field of the present observations have been sketched in Figure 1. Certain sections of the H II region can definitely be associated with features of the radio contours.

The principal part of the nebula surrounds γ Velorum at $l = 230^\circ$, $b = -7^\circ$. This is clearly visible on the 3.5 m isophotes as a spur, several degrees wide, extending southwards from the belt of emission along the galactic equator. The excess equivalent aerial temperature at the centre of this object is about 500 °K, and its integrated flux density of the order of 500×10^{-26} m.k.s. units; an accurate determination is hindered by the proximity of the intense extended

source Vela-X, which is considered to be unrelated to the H II region. This H II region is of large angular size, and so, by Section III, the emission measure is $2 \times \Delta T_a$ or 1000; rather less than the 3000 estimated optically by Gum for the brighter parts. The difference may be partly due to the "patchiness" visible in photographs of the nebula, since the optical figure refers to the brightest portions, the radio measurement to a smoothed average.

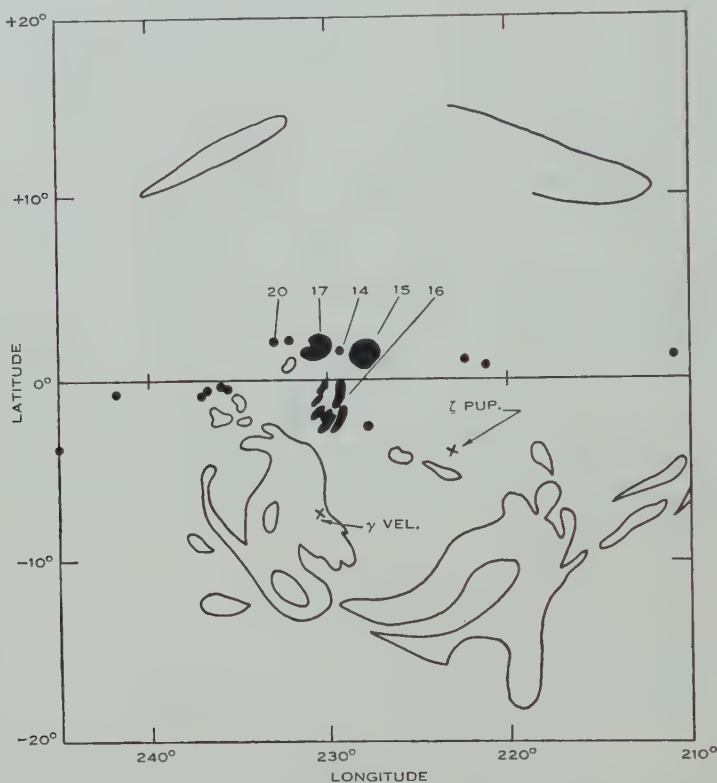


Fig. 3.—Sketch of H II regions (after Gum 1956). The large, unshaded regions are the several parts of Stromlo 12. The other H II regions referred to in the text (Stromlo 14, 15, 16, 17, and 20) are indicated by their numbers.

At 15 m a large area of absorption agrees well in position with the H II region. At the deepest point, listed in Table 2 as *R 104*, the absorption is about 60 per cent., consistent with $\tau=1$ or $E \approx 800$ by equation (2), but the optical emission measure of 3000 suggests that the region should absorb 98 per cent. of the background 15 m radiation. The excess of T_a over the electron temperature, about 25,000 °K, would then have to be ascribed to radiation originating between the nebula and the Sun.

The other principal section of the H II region, near $l=220^\circ$, $b=-10^\circ$, is not plainly visible at 3.5 m; it lies in an area of pronounced gradient of brightness temperature. A value $\Delta T_a \approx 200^\circ \text{K}$ can tentatively be assigned, leading to

a value of $E \approx 400$. The form of the contours suggests that the ridge of emission just to the north, listed as source R 34, may be associated with the nebula (its neighbourhood seems somewhat obscured optically). At 15 m there is clear evidence of absorption of at least 30 per cent. around longitude 220° , extending from $b = -9^\circ$ to $b = -16^\circ$: this information is gained from a recent record commencing at -18° . R 100 is another piece of the H II region. The extensive 15 m source R 102, which may be a blend of three distinct peaks of brightness, could hardly be seen through a moderately intense H II region, and so would either lie in front of the latter or be seen through a "window"; it might well be the 3.5 m source R 34, which would thus be non-thermal in character. The same applies to R 103, which lies on the boundary of the H II region at 231° , -11° and which may be identical with R 44.

Some other fragments of the nebula Stromlo 12 can be found on the radio isophotes. The arm at 235° , -8° appears at 3.5 m ($E=400$), although it is merged with another, probably unconnected, object—the source R 52—and there is a suggestion of its presence at 15 m. The outlying portion Stromlo 12b is identified with the extended object R 76 ($E=600$) and may be present at 15 m. A patch of emission at 218° , $+12^\circ$ is outside the area covered by the isophotes, but examination of records of the neighbourhood showed the presence of weak 3.5 m emission, with $E \sim 250$.

So far, the emphasis has been on searching the radio isophotes for features which correspond to definite patches of optical H II emission. However, there are not only patches of emission which have no marked effects on the isophotes (for example, near 225° , -12°) but also radio features where no H II emission has been observed optically (for example, the well-marked absorption at 15 m, R 114, which coincides with an extended 3.5 m source, R 73, at 235° , $+10^\circ$). These apparent inconsistencies are probably associated with the effects of interstellar absorption on the optical observations.

The possibility of observing, at radio wavelengths, an optically obscured H II region is obvious. For the rest, it should be remembered that in the above discussion only the well-marked radio features have been considered. It is possible that H II emission actually exists over a considerable area of sky in this region, so that some part of the radiation attributed to the background is, in fact, due to the extended H II region. This would explain to some extent the comparatively low values of emission measure deduced from the 3.5 m observations, since the values quoted would correspond to the excess of the brighter features over the H II background.

Quite distinct from the large Stromlo 12, a number of smaller H II regions lie near sources listed in Tables 1 and 2.

Stromlo 14 ($227^\circ.9$, $+1^\circ.4$)

This object is of appreciable angular diameter ($100'$), but it is listed as "very faint" by Gum. The 3.5 m source R 57 lies within a degree of its centre, and if the identification is accepted the emission measure is about 500. There is a trace of absorption at 15 m, although this is not included in Table 2.

Stromlo 15 ($229^{\circ}.2$, $+1^{\circ}.6$)

Although "moderately bright", the small diameter ($20'$) of this nebula would make it a faint object at 3.5 m. It does, however, coincide with a small spur on the contour map, and with $\Delta T_a = 100^{\circ}\text{K}$, as is possible, E could be as high as 2000 .

Stromlo 17 ($230^{\circ}.6$, $+1^{\circ}.9$)

This nebula, being "moderately bright" and 100 by $65'$ in size, might be thought to offer a good chance for detection. Unfortunately, its situation on the flank of an intense non-thermal source renders it quite invisible at 3.5 m. At 15 m, *R 109* agrees in position with this nebula; the observed absorption of about 20 per cent. implies an emission measure greater than 600 , the actual value depending on the amount of foreground emission.

Stromlo 20 ($232^{\circ}.9$, $+2^{\circ}.1$)

This H II region lies within $\frac{1}{2}^{\circ}$ of the source *R 62*, but an identification is most unlikely. In the first place, the position of the source has a probable error of only about $0^{\circ}.1$, and in any case, the small diameter ($7'$) of the H II region means that it would be invisible at 3.5 or 15 m unless it were exceedingly bright, or unless the radio emission were non-thermal.

From the results presented above, particularly the study of *Stromlo 12*, it is clear that, although the values of emission measure deduced from the radio observations are rather low, the radio emission from the H II regions can be adequately accounted for in terms of thermal emission and absorption.

(b) Non-thermal Emission and the Galactic Background

The most conspicuous radio features of the Vela-Puppis region are the intense extended sources in the galactic plane. Three peaks are much more intense than the others and form a close group (*R 56*, *R 58*, and *R 59*, designated Vela-X, -Y, and -Z respectively). They are not far from the well-known source Puppis-A, which, together with the others, is indicated in Figure 1. At 15 m Vela-Y and -Z are not separated and only two peaks are listed in Table 2 (*R 107* and *R 108*). Their non-thermal nature is shown by the very high equivalent aerial temperatures—more than $200,000^{\circ}\text{K}$ —recorded at 15 m. Away from these peaks, the brightness temperature along the equator is about $100,000^{\circ}\text{K}$ at 15 m, 3000°K at 3.5 m, and (from Piddington and Trent 1956*b*) about 15°K at 0.5 m; that is, roughly proportional to the $5/2$ power of the wavelength. The half-intensity width of the equatorial emission is about 6° , but comparison of the isophotes at the three wavelengths suggests that this width decreases slightly with increasing wavelength.

The 3.5 m contours show a line of extended sources along the galactic plane. Vela-Y and -Z lie in the direction of the Vela dark nebula (Greenstein 1937), and may represent regions of interaction between interstellar dust, fast particles, and magnetic fields, in accordance with current theories of non-thermal radio emission. The other sources along the equator may be similar objects.

Apart from these sources, there is a gradual increase of 3.5 m brightness temperature along the galactic equator from longitude 220° to 240° . A similar

effect, complicated by equatorial absorption (*R 111*) between 235 and 240° , occurs at 15 m and this may be related to the spiral structure of the Galaxy. Vela-Puppis lies about 90° from the galactic centre and the path length through the local arm which is included in the line of sight increases with increasing longitude (see, for example, Fig. 1 of Kerr, Hindman, and Carpenter 1957). If the emitting regions are optically thin, then the brightness temperatures depend on the total depth of the emitting material in the line of sight, so that, at least in this limited region of the Galaxy, the metre wavelength observations are consistent with the origin of non-thermal radiation within spiral arms.

It is of interest to note that the 3.5 m brightness temperature in the regions about 10° distant from the galactic equator on the southern side (except around large H II regions) is smaller than that at an equal distance north of the equator. Between latitude $+10^\circ$ and the limit of the survey at about $+13^\circ$ the measured brightness temperature never falls to 1500°K , whereas much of the zone between -10 and -13° is near 1250°K . Further from the galactic plane, at -20 to -25° , the brightness temperature is uniform over large tracts of sky at 1100°K —little more than the values of about 900° observed at this wavelength in the coldest parts of the sky.

(c) *Discrete Sources*

The 77 sources in the 3.5 m list show no peculiarities of distribution. The density in the vicinity of latitudes -10 to -15° is rather greater than the average for the whole region surveyed of 0.10 source per square degree, but the difference does not seem to be significant. Many of the objects may be extragalactic.

A comparison of detailed surveys at different wavelengths made with similar instruments has seldom been possible hitherto, and so it is worth considering together Tables 1 and 2. The number of coincidences within 1° is nine, whilst about four would be expected to occur by chance. Very strong sources such as Puppis-A, Vela-X, and Vela-Y are certainly observed at both wavelengths.

Only four of the other sources seem to have been reported by previous observers (see Table 1 for details). Optical identifications seem to be out of the question for the vast majority of the sources; only three close coincidences between Table 1 sources and NGC or IC objects have been noted. The optical objects concerned—NGC 2220, 2477, and 2547—are star clusters and there is no apparent reason why they should be radio emitters.

The source designated R 54 and R 105 is the well-known Puppis-A (IAU 08S4A), identified with a peculiar galactic nebulosity and probably rather similar to the source Cassiopeia-A. A detailed study of the source at 3.5 m wavelength has been made by Sheridan (1958), and the flux density quoted in Table 1, which is the author's estimate, must be regarded as provisional. Values of its flux density at several other wavelengths have been reported, but it is now clear that most of the observations were subject to confusion with the other intense sources nearby. It is therefore profitless to derive a detailed spectrum from existing data. Qualitatively, however, it appears that at the higher frequencies the flux density increases with wavelength, but there is little change

between 3.5 and 15 m. How much of this flattening is due to absorption in interstellar H II regions is not known.

Another interesting source is Vela-X (R 56, R 107), the brightest 3.5 m object. Making an approximate allowance for the width of the aerial beam, the angular size of the source to half-brightness is about $1\frac{3}{4}^\circ$ in Right Ascension and about $1\frac{1}{4}^\circ$ in declination. Gum (personal communication) has suggested its identification with Stromlo 16, an unusual filamentary nebulosity reported by Melotte (1926). However, the nebula is about 4 by 2° in angular size, larger than the 3.5 m source, and the most prominent filaments lie at 230° , -2° , where the radio emission is comparatively weak. The total 3.5 m flux density is greater than that of IC 443, but the latter is brighter optically than Stromlo 16 and appears to show a correlation between optical and radio features (Rishbeth 1956). Thus the suggested identification of Vela-X remains uncertain.

The extended source in Vela-Puppis observed at 75 cm wavelength by McGee, Slee, and Stanley (1955), and at 50 cm by Piddington and Trent (1956*a*), is a blend of Vela-X, -Y, -Z and galactic background. However, more detailed analysis suggests that at these wavelengths Vela-X is the outstanding object in this vicinity, and indicates further that the spectrum of this source may be almost flat in the range of wavelengths for which data are available.

By far the strongest source away from the equatorial belt is R 8, with an integrated flux density of 150×10^{-26} m.k.s. units, or almost 200×10^{-26} units if combined with the weaker source R 6 nearby. Optical identification of such a prominent and isolated feature might be possible, but would perhaps be hindered by the proximity of the star Canopus.

V. CONCLUSION

It has been shown that the principal H II regions in Vela-Puppis can be detected in emission at 3.5 m wavelength and in absorption at 15 m, and that qualitatively the observed radiation corresponds to the expected thermal emission of these objects. No real evidence for non-thermal emission from H II regions is found. Quantitatively, there is a discrepancy in the sense that the radio emission is too weak. This may well be due to selection of the optical data or other effects arising from non-uniformity of the H II regions; all the objects considered are very faint by optical standards.

The most prominent radio sources in the Vela-Puppis region are located near the galactic plane. They are non-thermal, and are a few degrees in diameter. The variation of radio emission along the galactic equator may be related to the spiral structure of the Galaxy. Many other discrete sources have been located, but their distribution shows no significant departure from uniformity.

VI. ACKNOWLEDGMENTS

The author is indebted to Dr. J. L. Pawsey, to Dr. B. Y. Mills, and in particular to Mr. C. A. Shain for continual help and guidance. The work described in the present paper was carried out during the tenure of a research studentship awarded jointly by the Commonwealth Scientific and Industrial Research Organization in Australia and the Department of Scientific and Industrial Research in Great Britain.

VII. REFERENCES

- BOLTON, J. G., STANLEY, G. J., and SLEE, O. B. (1954).—*Aust. J. Phys.* **7**: 110.
 GREENSTEIN, J. L. (1937).—*Harvard College Observatory Annals* **105**: 359.
 GUM, C. S. (1955).—*Mem. R. Astr. Soc.* **67**: 155.
 GUM, C. S. (1956).—*Observatory* **76**: 150.
 KERR, F. J., HINDMAN, J. V., and CARPENTER, MARTHA S. (1957).—*Nature* **180**: 677.
 MCGEE, R. X., SLEE, O. B., and STANLEY, G. J. (1955).—*Aust. J. Phys.* **8**: 347.
 MELOTTE, P. J. (1926).—*Mon. Not. R. Astr. Soc.* **86**: 636.
 MILLS, B. Y., LITTLE, A. G., and SHERIDAN, K. V. (1956).—*Aust. J. Phys.* **9**: 218.
 MILLS, B. Y., LITTLE, A. G., SHERIDAN, K. V., and SLEE, O. B. (1958).—*Proc. Inst. Radio Engrs.*, N.Y. **46**: 67.
 MILLS, B. Y., and SLEE, O. B. (1957).—*Aust. J. Phys.* **10**: 162.
 PIDDINGTON, J. H. (1951).—*Mon. Not. R. Astr. Soc.* **111**: 45.
 PIDDINGTON, J. H., and TRENT, G. H. (1956a).—*Aust. J. Phys.* **9**: 74.
 PIDDINGTON, J. H., and TRENT, G. H. (1956b).—*Aust. J. Phys.* **9**: 481.
 RISHBETH, H. (1956).—*Aust. J. Phys.* **9**: 494.
 SHAIN, C. A. (1958).—*Proc. Inst. Radio Engrs.*, N.Y. **46**: 85.
 SHAIN, C. A., and HIGGINS, C. S. (1954).—*Aust. J. Phys.* **7**: 130.
 SHERIDAN, K. V. (1958).—*Aust. J. Phys.* **11**: 400.

APPENDIX I

TABLE I

RADIO SOURCES OBSERVED AT 3.5 M WAVELENGTH IN THE VELA-PUPPIS REGION

No.	Galactic Ref.	South		Peak Temp. ΔT_a (°K)	Flux Density S (10^{-26} m.k.s.)	Notes*
		R.A. (1950) h m	Dec. ° '			
R 1	213—25	05 59.5	39 30	300	—	Ridge 3°
R 2	225—27	06 06.0	50 00	275	—	Ext.
R 3	222—25	06 12.3	47 20	450	28	(Puppis-A, Pictor-A)
R 4	212—21	06 17.7	37 05	400	24	
R 5	220—23	06 19.8	45 00	275	17	(Fornax-A) (NGC 2220)
R 6	228—25	06 20.4	52 40	700	44	Part of R 8 ?
R 7	211—19	06 25.1	35 25	450	28	S06-3
R 8	229—24	06 25.3	53 38	2150	150	S06-5
R 9	216—20	06 30	40	150	—	Ridge 3°
R 10	212—18	06 31.0	36 40	300	—	Ext.
R 11	227—22	06 38.0	51 00	225	14	
R 12	213—16	06 41.0	36 10	175	11	
R 13	217—17	06 46.4	39 45	575	34	
R 14	233—22	06 50.0	56 00	150	10	
R 15	217—15	06 56.0	39 30	175	11	
R 16	225—17	07 00.9	47 15	300	19	(Fornax-A, Pictor-A)
R 17	223—16	07 02.2	45 30	275	—	Ext.
R 18	221—15	07 04.3	42 45	300	19	
R 19	228—18	07 05.2	50 20	275	17	

* BSS: Bolton, Stanley, and Slee (1954); P: Piddington and Trent (1956a); S: Shain and Higgins (1954); Str: Gum (1955).

TABLE 1 (*Continued*)

No.	Galactic Ref.	R.A. (1950) h m	South Dec. ° '	Peak Temp. ΔT_a (°K)	Flux Density S (10^{-26} m.k.s.)	Notes*
R 20	218-13	07 06	40	300	—	Ridge 2°
R 21	215-11	07 07.2	35 45	300	18	
R 22	231-18	07 08.0	53 00	150	—	Ext.
R 23	230-17	07 10.5	51 30	125	8	
R 24	216-10	07 16	36 30	350	—	Ridge 2° : BSS 103
R 25	220-12	07 16.0	40 50	250	15	
R 26	234-18	07 18.6	55 30	300	19	
R 27	215-09	07 19.0	35 00	150	10	
R 28	232-16	07 20.3	52 50	350	—	Ext.
R 29	220-11	07 21.8	40 00	250	15	
R 30	216-08	07 26.0	35 40	300	18	
R 31	232-15	07 29.3	52 40	300	19	
R 32	216-06	07 32.0	34 40	150	10	
R 33	229-13	07 35.3	49 00	250	15	(Pictor-A)
R 34	219-06	07 38	37	250	—	Ridge 4°
R 35	234-15	07 40	54 30	150	—	Ext.
R 36	234-14	07 41.0	53 20	200	12	
R 37	229-10	07 46.0	47 50	125	8	
R 38	223-07	07 48.0	40 40	250	—	Ext.
R 39	226-08	07 48.1	43 58	550	33	
R 40	227-09	07 48.5	45 30	225	14	(Vela-X)
R 41	221-05	07 49	38 30	250	—	Ridge 3° (NGC 2477)
R 42	218-03	07 50.0	35 00	225	—	Ext.
R 43	225-07	07 51.2	43 00	250	15	(Puppis-A)
R 44	231-10	07 53.0	49 30	250	—	Ext.
R 45	237-14	07 54.0	56 10	150	10	
R 46	236-13	07 56.0	55 00	150	10	
R 47	230-08	07 59.5	47 20	225	14	
R 48	223-02	08 08.0	39 00	500	—	Ext.
R 49	232-08	08 09.4	49 30	600	38	(NGC 2547)
R 50	227-04	08 09.8	42 40	275	17	(Puppis-A)
R 51	222-00	08 12	36	250	—	Ext.
R 52	236-10	08 13	54	250	—	Ext.
R 53	230-06	08 13.5	46 40	325	20	
R 54	228-03	08 21.2	42 50	7000	750	08S4A, Puppis-A
R 55	227-00	08 26.0	40 20	300	—	Ext.
R 56	232-02	08 33.8	45 30	5600	1600	Vela-X
R 57	228+02	08 41.0	39 45	250	15	Str 14
R 58	231+00	08 42.5	43 30	2500	—	Vela-Y. Ext.
R 59	233-01	08 47.0	45 30	2500	—	Vela-Z. Ext.
R 60	225+07	08 49.0	34 40	225	14	
R 61	236-00	08 58.0	47 20	600	—	Ext.
R 62	233+02	08 59.3	43 20	300	19	(Str 20 ?)
R 63	227+08	09 00	35 30	400	—	Ext.—part R 64 ?

* BSS : Bolton, Stanley, and Slee (1954) ; P : Piddington and Trent (1956a) ; S : Shain and Higgins (1954) ; Str : Gum (1955).

TABLE 1 (Continued)

No.	Galactic Ref.	R.A. (1950) h m	South Dec. ° '	Peak Temp. ΔT_a (°K)	Flux Density S (10^{-26} m.k.s.)	Notes*
R 64	230+06	09 01	38 30	600	—	Ridge: BSS 119, P 17 ?
R 65	238-00	09 10	49 20	1200	—	Ext.
R 66	239-01	09 12	50 10	1000	—	Ext.—part R 65 ?
R 67	243-03	09 16.9	54 45	650	42	(Hya-A, Centaurus-A)
R 68	235+06	09 23.3	42 00	225	14	
R 69	241+01	09 25.8	50 20	275	17	
R 70	238+05	09 28.5	45 20	150	10	
R 71	241+02	09 34.0	49 10	250	15	
R 72	239+05	09 35.5	45 30	175	11	
R 73	235+10	09 38	39	150	—	Ridge 2°
R 74	243+02	09 42	50	350	—	Ext.
R 75	239+09	09 46.5	42 20	250	15	
R 76	236+12	09 48	38	275	—	Ridge 3° Str 12b
R 77	242+06	09 49.0	46 50	250	—	Ext.

* BBS: Bolton, Stanley, and Slee (1954); P: Piddington and Trent (1956a); S: Shain and Higgins (1954); Str: Gum (1955).

TABLE 2
RADIO SOURCES OBSERVED AT 15.2 M WAVELENGTH

No.	Galactic Ref.	R.A. (1950) h m	South Dec. ° '	Peak Temp. (10^3 °K)	Flux Density (10^{-26} m.k.s.)	Notes*
R 100	227-12	07 31	47 00	-30	—	Part of Str 12a
R 101	225-11	07 33	45 00	40	320	
R 102	221-08	07 40	40	45	—	Complex (R 34)
R 103	231-11	07 50	49 30	25	200	(R 44)
R 104	230-07	08 08	47	-40	—	Ext.—Str 12a
R 105	228-03	08 21	43 00	100	800	08S4A, Puppis-A
R 106	235-04	08 37	49 30	30	240	
R 107	232-02	08 38	45 30	45	360	Vela-X (R 56)
R 108	231-00	08 44	44 00	~75	~600	Vela-Y+Z (R 58+R 59)
R 109	231+02	08 50	42 00	-30	—	Str 17
R 110	239-02	09 03	51	40	320	Ext.
R 111	237+00	09 08	48	-35	—	Trough 4° (R 65)
R 112	237+04	09 22	45 30	30	240	
R 113	243+00	09 33	52	55	440	Ext. (R 69)
R 114	236+10	09 38	40 00	-25	—	(R 73)
R 115	242+06	09 54	47	45	360	Ext. (R 77)

* Str: Gum (1955).

RADIATION TRANSFER AND THE POSSIBILITY OF NEGATIVE ABSORPTION IN RADIO ASTRONOMY

By R. Q. TWISS*

[Manuscript received July 28, 1958]

Summary

Stimulated transitions are relatively enormously more probable at radio than at optical frequencies and it is this which makes it possible for negative absorption to arise at radio wavelengths when the medium will behave like an amplifier to the incident radiation. A necessary condition for the existence of this phenomenon is that the kinetic energy distribution $F(\eta)$ of the radiating electrons be markedly non-thermal with an appreciable excess of high energy electrons such that $\partial F/\partial \eta$ is positive over a finite range of the kinetic energy η . However, this condition is not sufficient, since it is shown that an electron gas in which free-free transitions provide the dominant radiation process can never exhibit negative absorption whatever the form of $F(\eta)$, and it is further necessary that the stimulated transition probability should have a maximum at some finite value of the kinetic energy, the most favourable case occurring when this maximum is a sharp one at the value of η at which $\partial F/\partial \eta$ has a positive maximum. These conditions can both be met in principle for the cases in which the dominant radiation process is due (a) to Cerenkov effect, (b) to gyro radiation by non-relativistic electrons, (c) to synchrotron-type radiation by highly relativistic electrons, and it is shown that negative absorption can arise in all these cases; the relevance of these results to radio astronomy is discussed briefly.

I. INTRODUCTION

The fundamental equations governing the transfer of radiation in an ionized medium are of the same general form at both radio and optical wavelengths. However, the physical nature of the solution in a given case depends upon the specific processes governing the emission, absorption, and scattering of radiations; since these are very different at optical and at radio wavelengths, one may expect to find corresponding differences between the characteristics of radiation transfer.

Thus at optical wavelengths the transfer of radiation is appreciably affected by the radiative transitions between bound states which give rise to the characteristic spectral lines. However, at radio wavelengths the number of discrete lines such as the 1420 Mc/s hyperfine line of atomic hydrogen is extremely limited. Furthermore, the few lines that do exist are likely only to be observable in interstellar gas clouds in which the density is so low that the refractive index differs negligibly from unity. Under these conditions the effect of these spectral lines on the transfer of radiation is an extremely simple case of the effect encountered at optical wavelengths and will not be discussed here. Instead, we shall deal only with the case where radiation is emitted or absorbed in free

* Division of Radiophysics, C.S.I.R.O., University Grounds, Chippendale, N.S.W.

electron transitions in which the available electron energy is much greater than $h\nu$, the energy carried by a single quantum of radiation. We shall further assume that the radiation is incoherent in origin in the sense that the probability of a radiative transition may depend upon the *average* density of matter and radiation but not upon any time-dependent variation in these. Hence we shall not consider the case where the radiation is produced by plasma oscillations, a coherent process, but only processes such as those involved in free-free transitions, Cerenkov radiation, gyro radiation, and synchrotron radiation.*

II. COMPARISON BETWEEN THE OPTICAL AND RADIO CASES

If the density of a radiation field at a frequency ν is equal to that of a black-body radiation field at temperature Θ_ν , then, in a typical stellar atmosphere, the dimensionless ratio $k\Theta_\nu/h\nu$ lies usually between 1 and 0.1 at optical wavelengths. However, at a frequency of 10^8 c/s, which lies in the metre wave radio band, $k\Theta_\nu/h\nu \sim 10^8$ for the quiet Sun and may be as high as 10^{14} in the neighbourhood of a solar burst; similarly, near the intense radio sources Cassiopeia-A and Cygnus-A, $k\Theta_\nu/h\nu \sim 10^8$ and 10^9 respectively for $\nu \sim 10^8$ c/s, and even in interstellar space, not in the immediate neighbourhood of a radio source, $k\Theta_\nu/h\nu \sim 10^4$ at this frequency. One consequence of this is that stimulated emission is relatively much more important at radio than at optical wavelengths, since, in the simple case in which the refractive index μ of the medium may be taken as unity, it follows immediately from the radiation laws† that γ , the ratio of the probability of stimulated to spontaneous emission, is given by

$$\gamma = [\exp(h\nu/k\Theta_\nu) - 1]^{-1}. \quad \dots\dots\dots (1)$$

At optical wavelengths it is usually the case that

$$\gamma \sim \exp(-h\nu/k\Theta_\nu) < 10^{-1},$$

but at radio wavelengths, where in general

$$\gamma \sim k\Theta_\nu/h\nu,$$

stimulated transitions can be enormously more probable than spontaneous transitions. In the non-equilibrium case, in which the density distribution over the accessible energy states of the matter field differs, by definition, from the thermal one, the relatively large probability for stimulated transitions may have important consequences, as we shall show below.

Equation (1), which is valid when $\mu=1$, may reasonably be expected to apply under the conditions prevalent in the discrete radio sources or in interstellar space, but not under the conditions prevalent in a stellar corona. In the latter case, not only will μ be different from unity but, in the presence of magnetic fields such as those associated with sunspots, the medium will be birefringent, in strong contrast to the conditions at optical wavelengths when it is valid to take $\mu=1$ even in the stellar photosphere.

* By gyro radiation we mean radiation at the fundamental or at the first few harmonics of the gyro frequency by weakly relativistic electrons rotating in a magnetic field. By synchrotron radiation we mean radiation by strongly relativistic electrons at high harmonics of the gyro frequency.

† See e.g. Fowler (1936), Chapter XIX.

A further characteristic of the radio case is that the emission and absorption of radiation is often markedly anisotropic. This is especially the case for synchrotron radiation at or near the critical frequencies in a plasma permeated by a magnetic field; in the optical case, on the other hand, the emission at least can almost always be taken as isotropic.

Fortunately, these features in which the radio case is more complex than the optical are more than balanced by others for which the reverse is true. In the first place, the effects of incoherent microscopic scattering, which are so important in the transfer of optical radiation, can almost always be completely ignored in the radio case (Smerd, personal communication 1958), when appreciable scattering will only arise from irregular macroscopic variations in the electron density. Secondly, the kinetic energy of a typical particle of the matter field is so much greater than $h\nu$ that the emission or absorption of a radio-frequency quantum has a negligible effect upon the overall energy distribution of the matter field; a fact that provides the principal justification for the use of a classical radiation theory in calculating the transition probabilities at radio wavelengths.

III. EMISSION AND ABSORPTION COEFFICIENTS AT RADIO FREQUENCIES

When the refractive index of the medium μ_ν is not equal to unity the equation of radiation transfer can be written in the form (Woolley 1947)

$$\mu_\nu^2 d(I_{(\nu)}/\mu_\nu^2)/ds = j_\nu - \mathbf{K}_\nu I_{(\nu)}, \quad \dots\dots\dots (2)$$

where $I_{(\nu)}$ is the intensity of radiation in the direction \mathbf{s} and where j_ν , \mathbf{K}_ν are the volume coefficients of emission and absorption appropriate to the direction \mathbf{s} .

When the medium is birefringent a separate transfer equation must be set up for each mode. In the general case, in which macroscopic scattering must be allowed for, these two transfer equations will be coupled by terms in the absorption coefficient which represent scattering from one mode to the other. However, we shall ignore macroscopic scattering in what follows, and in this case, as has been pointed out by Martyn (1948), \mathbf{K}_ν simply represents the difference between the effects of stimulated absorption and stimulated emission in a single mode.

(a) *The Absorption Coefficient*

Accordingly, if n_{r_2} , n_{r_1} are the numbers of electrons in the r_2 and r_1 states respectively, if the energy difference between these states is $h\nu$, and if $B_{r_2,12}I(\nu)d\nu$, $B_{r_1,21}I(\nu)d\nu$ are the stimulated absorption and emission probabilities respectively for radiation with intensity $I(\nu)$ in unit solid angle in the direction \mathbf{s} , then

$$\mathbf{K}_\nu = \sum_r (n_{r_1} B_{r_1,12} - n_{r_2} B_{r_2,21}) h\nu. \quad \dots\dots\dots (3)$$

The Einstein coefficients $B_{r_2,12}$ and $B_{r_1,21}$ are not independent but are related by the equation

$$\tilde{\omega}_{r_2} B_{r_2,21} = \tilde{\omega}_{r_1} B_{r_1,12}, \quad \dots\dots\dots (4)$$

where $\tilde{\omega}_{r_2}$, $\tilde{\omega}_{r_1}$ are the statistical weights of the r_1 and r_2 states respectively.

In the optical case one can nearly always assume that some sort of local equilibrium has been reached and that the radiative transitions between states

r_1 and r_2 are in balance with the transitions induced by collisions or other means. This provides the justification for calling the stellar radiation at optical wavelengths "thermal" and a similar assumption is valid for the "thermal" radio sources such as the H II regions and the undisturbed solar corona. However, for the non-thermal radio sources this assumption is surely invalid since the radiation is produced in regions of extremely low density where a disturbed non-equilibrium state can exist for a comparatively long time. This is particularly the case for the discrete radio sources which, it is thought, are formed by clouds of very high energy electrons trapped by weak magnetic fields (Alfvén and Herlofson 1950). The lifetime of such an electron can be many thousands of years (Hoyle 1954, Twiss 1954) and it is extremely improbable that such a distribution of relativistic electrons can in any sense be regarded as in local equilibrium. Similarly, in the solar corona it is quite possible for a transient non-equilibrium electron energy distribution to last for times appreciably longer than the duration of the violent outbursts of radio noise which can often be measured in seconds (Pawsey and Bracewell 1955):

Under these circumstances the value of K_ν can be very different from that of the thermal case and it is at least conceivable that conditions could temporarily arise in which K_ν was negative, when the medium would behave like an amplifier. In the laboratory this condition can be achieved in the so-called maser (Gordon, Zeiger, and Townes 1955). The possibility that a similar effect could arise in nature is examined in Section IV below.

(b) *The Emission Coefficient*

If $A_{r,21} d\nu$ is the probability of a spontaneous transition with the emission of a quantum of given polarization into unit solid angle in the direction s , so that

$$j_\nu = \sum_r n_{r,2} A_{r,21} h\nu, \quad \dots\dots\dots (5)$$

then, in the case $\mu=1$, it is a well-known consequence of the radiation laws that

$$A_{r,21}/B_{r,21} = h\nu^3/c^2. \quad \dots\dots\dots (6)$$

However, when μ is different from unity this simple relation is no longer valid. Admittedly we know from Kirchhoff's laws that

$$\frac{j_\nu}{K_\nu} = \frac{\mu^2(\nu)\nu^2 h\nu}{c^2 [\exp(h\nu/k\Theta_\nu) - 1]}, \quad \dots\dots\dots (7)$$

when the matter field is in thermal equilibrium,* but we are not justified in concluding automatically from this that

$$A_{r,21}/B_{r,21} = \mu^2(\nu) h\nu^3/c^2, \quad \dots\dots\dots (8)$$

since, for one thing, $\mu(\nu)$ may depend upon the actual form of the energy distribution.

Accordingly, in the general case, which applies to the disturbed stellar corona, it is necessary to calculate the spontaneous and stimulated transition

* This ratio is smaller by a factor 2 than that normally quoted since we are dealing with each polarization separately in this paper.

probabilities separately and both of these may be appreciably different from their equilibrium values, especially at the critical levels at which $\mu^2(\nu)$ is near to infinity or to zero.

IV. CONDITIONS FOR NEGATIVE ABSORPTION

In the ideal case in which monochromatic radiation traversing a medium is only absorbed or emitted in transitions between a single pair of non-degenerate states, the necessary and sufficient conditions for negative absorption is that the higher energy states be more densely populated than the lower. However, in most cases of astronomical interest the initial and final states will possess both continuous and overlapping energy distributions, and in some cases these states will possess different degeneracies; under these conditions the simple criterion for negative absorption is no longer valid. Clearly a *necessary* condition is that at least some of the higher energy states be over-populated, but it may be that negative absorption will not arise however widely the distribution may depart from that of thermal equilibrium. One example of this is provided by a fully ionized medium in which radiation is assumed to be emitted and absorbed only by free-free transitions; this and other cases will be considered in detail below, but we shall first consider the problem in general terms, without specific reference to the mechanism of radiation, for the idealized case in which the energy of a state is the only dynamic variable that need be considered specifically. With this last simplification we may write equation (3) in the form

$$\mathbf{K}_\nu = \hbar\nu \int_0^\infty [n(\eta - \hbar\nu)B_{12}(\nu, \eta) - n(\eta)B_{21}(\nu, \eta)]d\eta, \quad \dots\dots (9)$$

where η is the energy of the state.

Let us now put

$$\left. \begin{aligned} F(\eta) &= n(\eta)/\tilde{\omega}(\eta), \\ Q(\eta) &= \tilde{\omega}(\eta)B_{21}(\eta), \end{aligned} \right\} \quad \dots\dots\dots (10)$$

where $\tilde{\omega}(\eta)$ is the statistical weight of the state of energy η . Then from equation (4) we can write

$$\mathbf{K}_\nu = -\hbar\nu \int_0^\infty \hbar\nu \frac{\partial F}{\partial \eta} Q(\eta) d\eta, \quad \dots\dots\dots (11)$$

since, in any case of radio-astronomical interest, $\eta \gg \hbar\nu$ for all η for which $Q(\eta)$ differs appreciably from zero.

The quantities $F(\eta)$, $Q(\eta)$ may be regarded as the effective energy distribution function and the effective cross section for stimulated transitions respectively. By definition both of these quantities are non-negative, that is,

$$F(\eta) \geq 0, \quad Q(\eta) \geq 0, \quad \dots\dots\dots (12)$$

while $F(\eta)$ must also satisfy the boundary condition

$$F(\infty) = 0, \quad \dots\dots\dots (13)$$

which is imposed by the requirement that no physically realizable state can have infinite energy.

If $F(\eta)$, $Q(\eta)$ are sufficiently well-behaved functions we can integrate equation (11) by parts to give

$$\mathbf{K}_\nu = (h\nu)^2 \left\{ F(0)Q(0) + \int_0^\infty F(\eta) \frac{\partial Q}{\partial \eta} d\eta \right\}. \quad \dots\dots\dots (14)$$

From equation (14) and the inequality (12) it follows immediately that \mathbf{K}_ν is positive, no matter what the form of $F(\eta)$, if $Q(\eta)$ is a monotonically *increasing* function of energy; similarly, it follows from equation (11) that \mathbf{K}_ν is positive no matter what the form of $Q(\eta)$ if $F(\eta)$ is a monotonically *decreasing* function of η .

However, if $Q(\eta)$ *decreases* monotonically with η , then \mathbf{K}_ν will be negative if the energy distribution function $F(\eta)$ has a large enough peak at some value greater than zero or if $F(0)=0$.

A case of particular interest is that in which $Q(\eta)$ has a strong resonance such that it is only appreciably different from zero over a small range of η centred in $\eta=\eta_0$. Under these circumstances it follows from equation (11) that \mathbf{K}_ν will be negative if

$$(\partial F / \partial \eta)_{\eta=\eta_0} > 0.$$

These general results will now be applied to a number of specific cases.

(a) *Free-free Transitions*

In the classical Lorentz theory of propagation in an ionized medium the effect of the collisions of the free electrons with the ions and neutral molecules is allowed for by introducing a damping term,

$$m\mathbf{v}_1\nu(\mathbf{v}_0),$$

into the equations of motion of an electron, where \mathbf{v}_1 is the velocity induced by the electromagnetic field and \mathbf{v}_0 is the undisturbed electron velocity. The quantity $\nu(\mathbf{v}_0)$, the collision frequency appropriate to an electron with velocity \mathbf{v}_0 which is calculated from kinetic theory, is inherently positive so that the effect of collisions is always to *abstract* energy from the electromagnetic field at a rate proportional to

$$\frac{1}{2}m \int \nu(\mathbf{v}_0) \mathbf{v}_1^2(\mathbf{v}_0) n(\mathbf{v}_0) d\mathbf{v}_0,$$

where $n(\mathbf{v}_0)d\mathbf{v}_0$ is the contribution to the electron density from electrons with velocities in the range \mathbf{v}_0 to $\mathbf{v}_0+d\mathbf{v}_0$. Hence on the *macroscopic* theory free-free transitions always involve a positive absorption of energy from the electromagnetic field, whatever the actual form of the electron velocity distribution function.

This conclusion can also be reached by considering the microscopic absorption and emission processes in the ionized gas as follows. We will restrict ourselves to the idealized case in which the ions can be taken as infinitely massive, a simplification that does not involve any significant loss of generality; we will also assume that the medium is not acted upon by an external magnetic field.

The theory of free-free transitions has been extended to a medium with refractive index $\mu(\nu)$ by Smerd and Westfold (1949), who showed that the spontaneous transition probabilities must be multiplied by a factor $\mu(\nu)$, while the stimulated transition probabilities must be multiplied by a factor $\mu^{-1}(\nu)$ if results consistent with thermodynamics are to be obtained. In this treatment it was assumed that the transition probabilities for a given electron are *only* affected by the presence of the other electrons in so far as these contribute to $\mu(\nu)$, a purely macroscopic effect, and the same assumption will be made here.

Let us assume that the number of electrons per unit volume with kinetic energies in the range η to $\eta + d\eta$ is given by

$$N_0 G(\eta) \eta^{\frac{1}{2}} d\eta,$$

where N_0 is the total number of electrons per unit volume. In the case where the distribution is Maxwellian

$$G(\eta) = \frac{2\pi \exp(-\eta/k\Theta)}{(\pi k\Theta)^{3/2}}, \quad \dots\dots\dots (15)$$

where Θ is the equilibrium temperature.

Besides the energy η the orbit of an electron with respect to a given scattering centre will be characterized by the impact parameter b , where b is the perpendicular distance between the scattering centre and an asymptote to the orbit. Then, following Fowler (loc. cit.), it can be shown that the number of encounters per scattering centre, in unit time, with (η_1, b_1) electrons is

$$N_0 2\pi b_1 db_1 (2/m)^{\frac{1}{2}} G(\eta_1) \eta_1 d\eta_1,$$

if we make the physically plausible assumption that the distribution over the accessible impact parameter states is purely random.

Now the probability of a transition from an (η_1, b_1) electron to an (η_2, b_2) electron with the absorption of a quantum of energy $h\nu$, where

$$h\nu = \eta_2 - \eta_1$$

may be taken to be

$$\beta_{12}(\eta, b) I(\nu) d\nu,$$

so that the number of transitions of (η_1, b_1) electrons to (η_2, b_2) electrons per scattering centre per unit time with absorption of a quantum of energy $h\nu$ is

$$N_0 (2/m)^{\frac{1}{2}} G(\eta_1) \eta_1 d\eta_1 2\pi b_1 db_1 \beta_{12}(\eta, b) I(\nu) d\nu.$$

Similarly the number of transitions of (η_2, b_2) electrons to (η_1, b_1) electrons with the emission of a quantum of energy $h\nu$ may be written

$$N_0 (2/m)^{\frac{1}{2}} G(\eta_2) \eta_2 d\eta_2 2\pi b_2 db_2 [\beta_{21}(\eta, b) I(\nu) d\nu + \alpha_{21}(\eta, b) d\nu],$$

where $\alpha_{21}(\eta, b)$ denotes the spontaneous transition probability.

To find the relation between β_{12} , β_{21} , and α_{21} we use the fact that detailed balancing must apply under the conditions of thermal equilibrium when $G(\eta)$ is given by equation (15) and $I(\nu)$ is given by

$$I(\nu) = \frac{\mu^2(\nu)\nu^2\hbar\nu}{c^2[\exp(\hbar\nu/k\Theta_\nu)-1]}, \quad \dots\dots\dots (16)$$

remembering that we are dealing with a polarized radiation field.

From this we obtain the relations

$$\eta_2 b_2 \beta_{21}(\eta, b) d\eta_2 db_2 = \eta_1 b_1 \beta_{12}(\eta, b) d\eta_1 db_1, \quad \dots\dots\dots (17)$$

and

$$\alpha_{21}(\eta, b)/\beta_{21}(\eta, b) = (\hbar\nu^3/c^2)\mu^2(\nu), \quad \dots\dots\dots (18)$$

so that equation (8) is valid for the case of free-free transitions.

From equation (17) it follows that the degeneracy of the state is proportional to the energy; a result that is to be expected when there is no preferred spatial direction.

If we substitute from equations (17), (18) in equation (3), replace sums by integrals, and integrate over the allowable values of η and b , we get

$$\mathbf{K}_\nu = \left(\frac{2}{m}\right)^{\frac{1}{2}} N_0^2 \hbar \nu 2\pi \int_0^\infty d\eta \int_0^{b_{\max}} db \eta b [G(\eta - \hbar\nu) - G(\eta)] \frac{c^2}{\hbar \nu^3 \mu^2(\nu)} \alpha_{21}(\eta, b), \quad \dots\dots\dots (19)$$

where b_{\max} is determined by the distance beyond which a given scattering centre is screened by the electronic cloud.

If the results given by Westfold (1950) are written in m.k.s. units, and in the form appropriate to a single state of polarization, we get

$$\alpha_{21}(\eta, b) = \frac{e^2}{3\pi^2 c^3 m \varepsilon_0} \frac{\eta \mu(\nu)}{1 + (8\pi \varepsilon_0 b \eta / Ze^2)^2} \frac{1}{\hbar \nu}, \quad \dots\dots\dots (20)$$

so that

$$\mathbf{K}_\nu = -\frac{1}{\nu^2 \mu(\nu)} \left(\frac{2}{m}\right)^{\frac{1}{2}} \frac{e^2 N_0^2}{3\pi c m \varepsilon_0} \left(\frac{Ze^2}{8\pi \varepsilon_0}\right)^2 \int_0^\infty d\eta \frac{\partial G}{\partial \eta} \ln \left[1 + \left(\frac{8\pi \varepsilon_0 b_{\max} \eta}{Ze^2}\right)^2\right], \quad \dots\dots\dots (21)$$

where Z is the ionic charge.

This equation is of the general form of equation (11) with

$$F(\eta) \sim G(\eta) \text{ and } Q(\eta) \sim \log [1 + (8\pi \varepsilon_0 b_{\max} \eta / Ze^2)^2], \quad \dots\dots (22)$$

so that in this case $Q(\eta)$ is a monotonically increasing function of η . It follows from the general discussion given above that \mathbf{K}_ν is positive, whatever the form of $G(\eta)$, in agreement with the predictions of the macroscopic theory. This conclusion, however, does not hold for other radiation mechanisms, as we shall now demonstrate.

(b) *Cerenkov Radiation*

The spontaneous emission probability for Cerenkov radiation is proportional to

$$\frac{e^2 v_0}{c} dv \left[1 - \frac{c^2}{v_0^2 \mu^2(v)} \right] \quad \text{for } v_0 > \frac{c}{\mu(v)}, \quad \dots \quad (23)$$

where v_0 is the velocity of the particle charge e traversing a medium with refractive index $\mu(v)$. In terms of the total energy E this may be written

$$P(E) \sim e^2 v dv \left(1 - \frac{m_0^2 c^4}{E^2} \right)^{\frac{1}{2}} \left[1 - \frac{E^2}{\mu^2(v)(E^2 - m_0^2 c^4)} \right] \quad \text{for } E > \frac{\mu(v) m_0 c^2}{\sqrt{\{\mu^2(v) - 1\}}} \\ = 0 \quad \text{for } E < \frac{\mu(v) m_0 c^2}{\sqrt{\{\mu^2(v) - 1\}}}, \quad \dots \quad (24)$$

which is zero below a certain critical value of the total energy E and thereafter increases monotonically with E . However, it does not follow from this and the general discussion given above that negative absorption is inherently impossible in this case, since Cerenkov radiation is not emitted isotropically but in a direction making an angle φ with the path of the charged particle where

$$\cos \varphi = c/v_0 \mu(v). \quad \dots \quad (25)$$

Hence, if we have a stream of fast particles moving, in a given direction, with a small spread in transverse velocities, the emission in a direction making an angle θ with the direction of the stream will be obtained from electrons with a narrow range of energies centred on $E = E(\theta)$, where

$$E(\theta) = \frac{m_0 c^2}{[1 - \sec^2 \theta / \mu^2(v)]^{\frac{1}{2}}}. \quad \dots \quad (26)$$

If the energy distribution function $G(E)$ is such that

$$(\partial G / \partial E)_{E=E(\theta)} > 0, \quad \dots \quad (27)$$

the absorption coefficient will be negative for propagation in this direction.

This condition is similar to that for growing longitudinal waves in an ionized plasma, which was first derived by Bohm and Gross (1949). These workers considered a unidimensional plasma with a velocity distribution function $H(v)$ and showed that a plasma wave with phase velocity v_0 would be amplified if

$$[\partial H(v) / \partial v]_{v=v_0} > 0. \quad \dots \quad (28)$$

However, there is one important difference between the two cases in that the growing plasma wave, which arises under conditions equivalent to a negative Landau damping (Landau 1946), is a coherent phenomenon while the negative Cerenkov absorption considered above is incoherent, so that, in this second case, the radiated power would tend to zero with the electronic charge e when the total current was kept constant.

(c) Gyro Radiation

If we consider an electron spiralling under the action of an external magnetic field with kinetic energy η then in the frame of reference of the rotating electron the frequency spectrum of the emitted radiation is confined to narrow bands centred on integral multiples of the gyro frequency

$$\nu_H = eB_0/2\pi m, \quad \dots\dots\dots (29)$$

where B_0 is the flux density of the magnetic field in weber/metre² and m is the relativistic transverse mass of the electron defined by

$$m = m_0(1 - v_0^2/c^2)^{-\frac{1}{2}}.$$

In the case where

$$m - m_0 \ll m_0, \quad \dots\dots\dots (30)$$

which we shall assume in the present section, appreciable radiation will only take place at the fundamental and at the first few harmonics of the gyro frequency. The width of these spectral lines measured in the frame of reference in which the axial velocity of the electrons is zero is determined essentially by the lifetime of an excited state, which is in turn affected by collisions with the background plasma ions and by diffusion into regions with slightly different magnetic fields. We shall assume that this fractional line width is very much less than the separation between successive harmonics, a condition which will certainly be valid in practice as long as the inequality (30) is satisfied. Under the conditions occurring in nature, the fast spiralling electrons will have different axial momenta so that the radiated frequency spectrum seen by a fixed observer will be Doppler broadened. However, in the present paper, in which we are primarily concerned to show that negative absorption is possible in principle, we shall only consider the idealized case in which the fast electrons all have zero axial momentum.

The presence of the external magnetic field removes the spin and angular momentum degeneracy of the electron, so that the weights of the states between which radiative transitions can take place are equal to unity.

Under these conditions we can apply the general results given in equations (9)-(14) above to write

$$K_\nu = -(\hbar\nu)^2 \int_0^\infty Q(\eta, \theta, \nu) \frac{\partial F}{\partial \eta} d\eta, \quad \dots\dots\dots (31)$$

where $F(\eta)$ is now simply the electron kinetic energy distribution function; $Q(\eta, \theta, \nu)$ is the stimulated transition probability in the direction θ for electrons of kinetic energy η .

When there is no Doppler broadening, $Q(\eta, \theta, \nu)$ is only appreciably different from zero for values of η close to the values

$$(\eta_r + m_0 c^2) = eB_0 c^2 / 2\pi r \nu, \quad r = 1, 2, \dots, \dots\dots\dots (32)$$

where r is the order of the harmonic of the gyro frequency so that we may write

$$Q(\eta, \theta, \nu) = \sum_{r=1}^\infty Q(\eta_r, \theta, \nu) \delta(\eta - \eta_r). \quad \dots\dots\dots (33)$$

If the energy of the fast electrons is sufficiently high, appreciable contributions to the power radiated at a given frequency ν can come from electrons with energies such that equation (32) is satisfied by one of several values of r . However, when the inequality (30) is valid, appreciable radiation at a given frequency will only occur for a single value of r , and in this case the associated value of \mathbf{K}_ν will be negative as long as

$$(\partial F / \partial \eta)_{\eta=\eta_r} > 0, \dots\dots\dots (34)$$

a condition that might arise in a practical case if a well-defined group of fast electrons was superimposed upon the thermal plasma electrons.

(d) *Synchrotron Radiation*

In this case where

$$m - m_0 \gg m_0,$$

the fast electrons are highly relativistic and appreciable energy is radiated over a very large number $\sim (m/m_0)^3$ of the harmonics of the gyro frequency. In the extreme case, which we shall consider here, the harmonics will be so closely spaced that the power spectrum $P(E, \varphi, \nu)$ for spontaneous emission may be assumed to be a smoothly continuous function given (Schwinger 1949) for the case $\mu=1$ to which we shall confine ourselves by

$$P(E, \varphi, \nu) = \frac{\sqrt{3} \mu_0 c}{2} \cdot \nu_{H0} \cdot \frac{\nu}{\nu_c} \int_{\nu/\nu_c}^{\infty} K_{5/3}(\zeta) d\zeta \text{ watts/cycle, } \dots (35)$$

where $K_\sigma(\zeta)$ is the modified Hankel function of order σ ,

$$\nu_c = 1.5 \nu_{H0} (E/m_0 c^2)^2 \sin \varphi,$$

$$\nu_{H0} = e B_0 / 2\pi m_0,$$

B_0 is the flux density in weber/m² of the external magnetic field.

φ is the angle between the direction of motion of the electron and the magnetic field.

The orbit of a given electron is characterized by four parameters :

E , the total energy ;

p_z , the axial momentum ;

r_0, ψ_0 , the cylindrical coordinates of the axis of the orbit.

In processes involving the emission or absorption of a quantum $h\nu$ in a direction making an angle θ to the magnetic field the changes δE and δp_z in the energy and axial momentum are given, from the conservation laws, by

$$\left. \begin{aligned} \delta E &= \pm h\nu, \\ \delta p_z &= \pm h\nu \cos \theta/c, \end{aligned} \right\} \dots\dots\dots (36)$$

the sign being taken positive for energy absorption and negative for energy emission. The coordinates of the axis of the orbit can also suffer changes which must be allowed for in a quantum calculation of the spontaneous emission probabilities (Judd *et al.* 1952 ; Olsen and Wergeland 1952), but there is no need to take explicit account of these in the present discussion since they have no influence on the form of the absorption coefficient.

The energy and momentum of the particle are related by the equation

$$cp_z = (E^2 - m_0^2 c^4)^{\frac{1}{2}} \cos \varphi. \quad \dots\dots\dots (37)$$

From this it can be shown that the change $\delta\varphi$ in the inclination of the electron orbit after absorption or emission of a quantum of energy $h\nu$ is given by

$$\delta\varphi \approx \pm \frac{h\nu}{\sin \varphi \sqrt{(E^2 - m_0^2 c^4)}} \left[\frac{E \cos \varphi}{\sqrt{(E^2 - m_0^2 c^4)}} - \cos \theta \right]. \quad \dots\dots (38)$$

Now let us assume that $N(E, \varphi)$, the combined distribution function for the total energy and orbital inclination of the relativistic electrons, has first-order derivatives with respect to both E and φ . Let

$$I(\nu, \theta) \beta_{12}(E, \varphi, \nu, \theta) \text{ and } I(\nu, \theta) \beta_{21}(E, \varphi, \nu, \theta) \quad \dots\dots (39)$$

be the stimulated absorption and emission probabilities for transitions between (E, φ) and $(E + \delta E, \varphi + \delta \varphi)$ states involving quanta of energy $h\nu$ emitted or absorbed in the θ -direction, where $I(\nu, \theta)$ is the intensity of radiation polarized so that the electric vector is perpendicular to the external magnetic field.* Then, from the fact that the electron states are non-degenerate, we have

$$\beta_{12}(E, \varphi, \nu, \theta) = \beta_{21}(E, \varphi, \nu, \theta), \quad \dots\dots\dots (40)$$

and

$$\mathbf{K}_\nu(\theta) = h\nu \int_{m_0 c^2}^{\infty} dE \int_0^\pi d\varphi \beta_{21}(E, \varphi, \nu, \theta) [N(E, \varphi) - N(E + \delta E, \varphi + \delta \varphi)]. \quad \dots\dots\dots (41)$$

Let $\alpha_{21}(E, \varphi, \nu, \theta)$ be the spontaneous emission probability in unit solid angle in the direction θ , so that

$$\alpha_{21}(E, \varphi, \nu, \theta) = \frac{P(E, \varphi, \nu)}{h\nu} \frac{G(\theta)}{2\pi \sin \theta}, \quad \dots\dots\dots (42)$$

where $G(\theta)$, the function which determines the angular distribution of the radiation, is subject to the relation

$$\int_0^\pi G(\theta) d\theta = 1. \quad \dots\dots\dots (43)$$

Since the refractive index μ of the medium has been assumed to be equal to unity, and since the radiation is emitted with a specific polarization, it follows from the basic radiation laws (Fowler loc. cit.) that

$$\beta_{21}(E, \varphi, \nu, \theta) = \frac{c^2}{h\nu^3} \alpha_{21}(E, \varphi, \nu, \theta). \quad \dots\dots\dots (44)$$

* This is the only polarization which interacts appreciably with relativistic electrons.

From equations (41)-(44) we see that

$$\begin{aligned} \mathbf{K}_\nu = -\frac{1}{2\pi \sin \theta} \int_{m_0 c^2}^{\infty} dE \int_0^{2\pi} d\varphi \frac{c^2}{\nu^2} P(E, \varphi, \nu) G(\theta) \left\{ \frac{\partial N}{\partial E}(E, \varphi) \right. \\ \left. + \frac{1}{\sin \varphi} \left(\frac{E \cos \varphi}{\sqrt{(E^2 - m_0^2 c^4)}} - \cos \theta \right) \frac{\partial N(E, \varphi)}{\partial \varphi} \right\}. \end{aligned} \quad (45)$$

Now the radiation from highly relativistic electrons is effectively confined to a narrow cone, with the axis tangential to the instantaneous orbit, so that $G(\theta)$ will only depart appreciably from zero when $\varphi \approx \theta$. In this case $\delta\varphi$, as given by equation (38), is effectively zero, as, therefore, is the coefficient of $\partial N(E, \varphi)/\partial \varphi$ in equation (45). If we make the further assumption that $\partial N(E, \varphi)/\partial E$ does not vary appreciably over the resonance of $P(E, \varphi, \nu)$ we may put

$$P(E, \varphi, \nu) G(\theta) \simeq P(E, \varphi, \nu) \delta(\theta - \varphi), \quad (46)$$

when from equation (45)

$$\mathbf{K}_\nu = \frac{-c^2}{2\pi \nu^2 \sin \theta} \int_{m_0 c^2}^{\infty} dE P(E, \theta, \nu) \frac{\partial N(E, \theta)}{\partial E}, \quad (47)$$

which is of the form of equation (11).

The function $P(E, \theta, \nu)$ is effectively zero for values of E such that

$$E \ll E_{\max.} \approx (\nu/\nu_{H0} \sin \theta)^{\frac{1}{2}} m_0 c^2, \quad (48)$$

risks to a maximum at $E_{\max.}$, and thereafter decreases monotonically to zero as E tends to infinity.

From the general discussion given above in the introduction to Section IV it follows that \mathbf{K}_ν is certainly negative, whatever the nature of $N(E, \theta)$ in the region $E > E_{\max.}$, as long as $N(E, \theta)$ is zero for all $E < E_{\max.}$. The most favourable case for negative absorption, for propagation in the θ -direction, arises when $N(E, \theta)$ is zero except for a narrow range of energies centred around the value at which $\partial P(E, \theta, \nu)/\partial E$ has a maximum, but this optimum is not critical since the maximum is not a sharp one.

Even if negative absorption is present in a medium it will only be important in regions in which the difference between the stimulated emission and absorption is at least comparable with the spontaneous emission. In the ideal case, where the relativistic electrons all have the same energy E , it can be shown quite simply that $\gamma(\theta)$, the ratio between the excess stimulated emission and the spontaneous emission, is given by

$$\gamma(\theta) = \frac{k\Theta_\nu(\theta)}{E} \frac{2\alpha(\partial X/\partial \alpha)}{X(\alpha)}, \quad (49)$$

where $\Theta_\nu(\theta)$ is the effective black-body temperature of the radiation field at frequency ν in the θ -direction, where α is given by

$$\alpha = \frac{2\nu}{3\nu_{H0} \sin \theta} \left(\frac{m_0 c^2}{E} \right)^2, \quad (50)$$

and where $X(\alpha)$ is the function defined by

$$X(\alpha) = \alpha \int_{\alpha}^{\infty} K_{5/3}(\zeta) d\zeta, \quad \dots\dots\dots (51)$$

which has been tabulated by Vladimírsky (1948).

The quantity

$$2\alpha(\partial X/\partial \alpha)/X(\alpha)$$

is a slowly varying function of α , for values of E at which appreciable radiation can be emitted at frequency ν , of the order of magnitude of unity so that even under ideal conditions appreciable negative absorption can only arise when

$$k\Theta_{\nu} \simeq \bar{E}, \quad \dots\dots\dots (52)$$

where \bar{E} is the average energy of the relativistic electrons which contribute appreciably to the radiation flux at frequency ν in the θ -direction.

V. DISCUSSION AND CONCLUSIONS

The main result of this paper has been to show that negative absorption or, in other words, amplification, can occur under certain circumstances in an electron gas.

A necessary condition for this is that the electron energy distribution function $F(\eta)$ be markedly non-thermal with an appreciable excess of high energy electrons such that the derivative of $F(\eta)$ is positive over a range of values of the kinetic energy η . However, this condition is not sufficient since, as we have shown, negative absorption can never occur if the dominant radiation process is that of free-free transitions, whatever the form of the electron energy distribution function. On the contrary, a further necessary condition is that the absorption cross section, in the direction of propagation, associated with the dominant radiation process should have a maximum at some finite value of the electron energy, and the most favourable case is when this maximum is a sharp one occurring at the value of η at which $\partial F/\partial \eta$ has its maximum positive value. More specifically, we have shown that negative absorption can occur in principle when the basic radiation process is (a) Cerenkov, (b) gyro, (c) synchrotron, and the relevance of this to radio astronomy will now be discussed briefly.

(a) Cerenkov Radiation from the Auroral Rays

A charged particle moving with velocity $c\beta$ through a medium will produce Cerenkov radiation in a particular mode and at given frequency ν in the direction θ for which

$$v_{\phi} = c\beta \cos \theta,$$

where v_{ϕ} is the phase velocity of the associated electromagnetic wave in the medium. Since, for a material particle, $\beta < 1$, this condition can be met only for $v_{\phi} < c$ and in an ionized medium this occurs only in the extraordinary mode at frequencies below the gyro frequency. Such radiation cannot escape from the Sun nor indeed from any galactic radio source (Pawsey and Bracewell 1955) at frequencies which can be transmitted through the Earth's ionosphere. It therefore appears that this effect cannot play any role in the production of radio

waves received from outside the Earth. However, it is conceivable that such radiation might be produced by fast electrons in the auroral rays (Ellis 1957) and, as long as the direction of these electrons is confined to a narrow solid angle, the discussion given in Section IV (b) suggests that the Cerenkov radiation generated in a certain solid angle will be amplified if the burst of fast electrons should have a sufficiently narrow energy spread.

(b) *Gyro Radiation and the Type I Solar Disturbances*

It still seems likely that the mechanism underlying the major outbursts of non-thermal radio emission from the solar corona is connected with the presence of organized plasma oscillations. However, it is possible that the strongly circularly polarized noise storms and solar bursts of spectral type I (Wild 1951) are due to harmonic gyro radiation (Twiss and Roberts 1958). If this is indeed the case one must assume the presence of negative absorption, or at least of anomalous radiation transfer, since for reasons connected with the circular polarization and lack of harmonic structure of the bursts the rotational energy of the fast gyrating electrons cannot be more than a few thousand electron-volts. This corresponds to a kinetic temperature of 10^7 – 10^8 °K, while the brightness temperature associated with the observed flux may be as high as 3×10^{10} °K. A discussion of this possibility must be based on a detailed analysis of the experimental evidence which will be published elsewhere (Twiss, in preparation); all we would like to emphasize here is that many of the objections to the gyro theory for type I disturbances can be resolved on the assumption that the medium is exhibiting negative absorption.

(c) *Synchrotron Radiation in the Discrete Radio Sources*

The principal objection to the theory which attributes the emission from the discrete radio sources to synchrotron type radiation is that one must assume that very large energies are stored both in the relativistic electrons and in the trapping magnetic field. These difficulties would be appreciably reduced if negative absorption were playing a significant role at least at frequencies below a few hundred megacycles. However, it seems virtually certain that this cannot be the case, as will now be shown.

Thus let us consider the brilliant discrete radio source Cygnus-A, which has an effective brightness temperature Θ_ν given approximately by the law

$$\Theta_\nu \simeq 10^9 (\nu/10^8)^{-3} \text{ °K}, \quad \dots\dots\dots (53)$$

over the frequency range

$$3 \times 10^7 < \nu < 10^{10} \text{ c/s.}$$

Now, as shown in Section IV, in the case most favourable for negative absorption, the energy of the fast electrons is given approximately by

$$E_{\text{opt}} \simeq m_0 c^2 \left(\frac{\nu}{1.5 \nu_{H0} \sin \theta \times 0.05} \right)^{\frac{1}{2}}, \quad \dots\dots\dots (54)$$

and for negative absorption to be significant it is then necessary that

$$E_{\text{opt}} \simeq k \Theta_\nu. \quad \dots\dots\dots (55)$$

The frequency at which this last condition is met is linearly proportional to $\nu_{H0}^{\frac{1}{2}}$, where $\nu_{H0} = eB_0/2\pi m_0$, is the gyro frequency of a non-relativistic electron in the magnetic trapping field. The average value of B_0 in Cygnus-A, which has been identified with two colliding galaxies, is probably within an order of magnitude of 10^{-8} weber/m² and it is almost certainly less than 10^{-6} weber/m², since otherwise the energy stored in the magnetic trapping field would be prohibitively large. Hence $\nu_{H0} < 10^5$ c/s and equation (55) will not be satisfied for frequencies above 2×10^7 c/s. Below this frequency the effective brightness temperature of Cygnus-A increases less rapidly than ν^{-3} (Lamden and Lovell 1956), while the opposite would be the case if negative absorption were starting to be effective. Admittedly, this effect might be partially marked by absorption in H II regions within our own galaxy, but, since we have taken very optimistic assumptions for the electron energy distribution and since Cygnus-A is an exceptionally brilliant source, the condition most favourable for stimulated emission, one is probably justified in concluding that negative absorption is always unimportant in the discrete radio sources, at frequencies of interest to radio astronomy, if the radio emission from these is indeed synchrotron-type radiation.

On the contrary, it is to be expected that the relativistic electron gas will exhibit positive absorption at low frequencies, as has been suggested earlier (Twiss 1954) and this phenomenon may be connected with the observed low-frequency cut-off in the spectrum of the discrete radio sources (Lamden and Lovell 1956).

VI. REFERENCES

- ALFVÉN, H., and HERLOFSEN, N. (1950).—*Phys. Rev.* **78** : 616.
 BOHM, D., and GROSS, E. P. (1949).—*Phys. Rev.* **75** : 1864.
 ELLIS, G. R. (1957).—*J. Atmos. Terr. Phys.* **10** : 302.
 FOWLER, R. H. (1936).—“Statistical Thermodynamics.” 2nd Ed. (Cambridge Univ. Press.)
 GORDON, J. P., ZEIGER, H. J., and TOWNES, C. H. (1955).—*Phys. Rev.* **99** : 1264.
 HOYLE, F. (1954).—*J. Geophys. Res.* **59** : 180.
 JUDD, D. L., LEFORE, J. V., RUDERMAN, M., and WOLFF, P. (1952).—*Phys. Rev.* **86** : 123.
 LAMDEN, R. J., and LOVELL, A. C. B. (1956).—*Phil. Mag.* **1** : 725.
 LANDAU, L. (1946).—*J. Phys. U.S.S.R.* **10** : 25.
 MARTYN, D. F. (1948).—*Proc. Roy. Soc. A* **193** : 44.
 OLSEN, H., and WERGELAND, H. (1952).—*Phys. Rev.* **86** : 123.
 PAWSEY, J. L., and BRACEWELL, R. N. (1955).—“Radio Astronomy.” (Oxford Univ. Press.)
 SCHWINGER, J. (1949).—*Phys. Rev.* **75** : 1912.
 SMERD, S. F., and WESTFOLD, K. C. (1949).—*Phil. Mag.* **40** : 831.
 TWISS, R. Q. (1954).—*Phil. Mag.* **45** : 249.
 TWISS, R. Q., and ROBERTS, J. A. (1958).—*Aust. J. Phys.* **11** : 424.
 VLADIMIRSKY, V. V. (1948).—*J. Exp. Theor. Phys.* **18** : 392.
 WESTFOLD, K. C. (1950).—*Phil. Mag.* **41** : 509.
 WILD, J. P. (1951).—*Aust. J. Sci. Res. A* **4** : 36.
 WOOLLEY, R. v. D. R. (1947).—Supplement to *Aust. J. Sci.* **10** : i.

SHORT COMMUNICATIONS

A SEARCH FOR RADIO EMISSION AT 3.5 M FROM THE LOCAL SUPERGALAXY*

By E. R. HILL†

As described by de Vaucouleurs (1953, 1956) the local supergalaxy is a large flattened system of galaxies and clusters of galaxies centred near the Virgo cluster. It is seen most clearly in the northern galactic hemisphere as a band of bright galaxies about 10° wide extending along a great circle. Our Galaxy appears to be near the rim of this system.

Radio emission apparently originating from the supergalaxy has been reported by Kraus and Ko (1953) and Kraus (1954) at 1.2 m and by Brown and Hazard (1953) at 1.9 m. In both instances the band of radio emission followed the band of galaxies quite closely; the combined observations extend about 90° along the supergalactic equator.

Particular interest has been added to radio observations of the local supergalaxy since the suggestion of Shklovsky (1954) that the emission arises from the synchrotron process which may be active in the intergalactic space within the supergalaxy. If this be so, radio observations might be expected to yield data relating to conditions in space between the galaxies—a region very difficult to study by any other means.

Observations to be described presently cast doubt upon the existence of radio emission from the supergalaxy. It should be added that other arguments leading to this negative conclusion have recently been reported by J. E. Baldwin and J. R. Shakeshaft at the I.A.U.-U.R.S.I. symposium on radio astronomy held at Paris in August this year.

Records taken with the Sydney 3.5 m cross-type aerial (Mills *et al.* 1958) have now been examined for traces of emission from the supergalaxy. The region studied in detail was from R.A. $10^h 30^m$ to $14^h 30^m$ between Dec. -15° and $+10^\circ$. Although this area has been investigated by Kraus and Ko, the high-resolution pencil beam of the "cross" (beamwidth $50'$ east-west by $70'$ north-south at these declinations) can provide a more detailed picture of the distribution of radiation. The results obtained, after a uniform gradient in temperature corresponding to the general fall-off in temperature away from the galactic plane has been removed, are shown in Figure 1. Effects of discrete sources have also been removed.

Also shown in Figure 1 are Shapley-Ames galaxies (dots) and the band, illustrated by the shading, observed by Kraus and Ko. The principal feature

* Manuscript received October 6, 1958.

† Division of Radiophysics, C.S.I.R.O., University Grounds, Chippendale, N.S.W.



Fig. 1.—The contours shown are of brightness temperature at 3.5 m after a uniform gradient stretching from R.A. 10^h 30^m, Dec. +10° to R.A. 14^h 30^m, Dec. -15° was removed. Contours are broken in areas subject to side-lobe effects of radio sources Virgo-A and Centaurus-A. Dots represent Shapley-Ames galaxies, and when circled represent two galaxies. The shaded region illustrates the band of emission observed by Kraus and Ko.

of the 3.5 m contours is the band of emission stretching from R.A. $13^{\text{h}} 06^{\text{m}}$ at Dec. $+10^\circ$ to about R.A. $11^{\text{h}} 32^{\text{m}}$ at Dec. $+15^\circ$, with a peak near R.A. $12^{\text{h}} 24^{\text{m}}$, Dec. 0° . It is clear that this feature does not conform to the band observed by Kraus and Ko nor does it appear to be related to the bright galaxies. Moreover, at the northern edge of the diagram at R.A. $12^{\text{h}} 28^{\text{m}}$ the aerial beam passes within 3° of the supergalactic centre and even here no excess emission is observed.

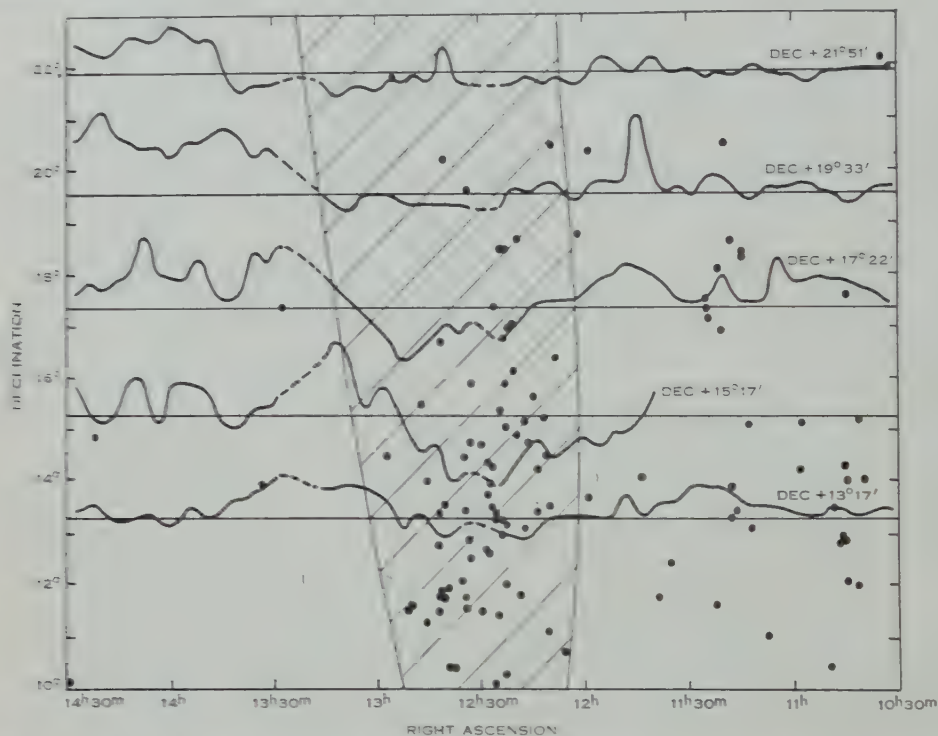


Fig. 2.—The radio band observed by Kraus and the distribution of bright galaxies between Dec. $+10^\circ$ and $+23^\circ$ are shown here in the same manner as in Figure 1. Superimposed are the additional records obtained with the Mills Cross at 3.5 m. The horizontal lines associated with each scan indicate the declination of the scan.

Additional observations at intervals of approximately 2° in declination have been made at 3.5 m from R.A. $10^{\text{h}} 30^{\text{m}}$ to $14^{\text{h}} 30^{\text{m}}$ between Dec. $+10^\circ$ and $+22^\circ$. Records so obtained are shown in Figure 2 superimposed on the band of emission observed by Kraus and the bright galaxies in the region. These 3.5 m observations carry less weight than those described above because of the larger zenith distances involved. They do, however, indicate that the feature noted at lower declinations retains its predominance and continues northward towards Dec. $+22^\circ$ at R.A. 14^{h} , and up to this declination continues to show no relation with the bright galaxies. Observations by Blythe (1957) at 8 m also show this feature, and comparison with the present results indicates that the radiation is non-thermal. From the extensive map given by Blythe, it seems possible that the

feature is an extremity of the well-known flare emanating from the galactic plane near $l=0^\circ$.

It is clear that between Dec. -15° and $+22^\circ$ emission from the supergalaxy is not apparent at 3.5 m in regions where earlier observations might have led one to expect it. Thus, if radio emission from the supergalaxy does exist in the region discussed, it is overshadowed by the feature described above.

References

- BLYTHE, J. H. (1957).—*Mon. Not. R. Astr. Soc.* **117**: 653.
 BROWN, R. H., and HAZARD, C. (1953).—*Nature* **172**: 997.
 KRAUS, J. D. (1954).—*Astrophys. J.* **59**: 113.
 KRAUS, J. D., and KO, H. C. (1953).—*Nature* **172**: 538.
 MILLS, B. Y., LITTLE, A. G., SHERIDAN, K. V., and SLEE, O. B. (1958).—*Proc. Inst. Radio Engrs.*, N.Y. **46**: 67.
 SHKLOVSKY, I. S. (1954).—*Astr. J. Moscow* **31**: 533.
 DE VAUCOULEURS, G. (1953).—*Astrophys. J.* **58**: 30.
 DE VAUCOULEURS, G. (1956).—"Vistas in Astronomy." (Ed. A. Beer.) Vol. 2, p. 1584. (Pergamon Press: London.)

THE CLOCK PARADOX IN SPECIAL RELATIVITY*

By H. JEFFREYS†

Dr. G. Builder (1957) has produced a new analysis of this problem, which is criticized by Professor Dingle (1957). I think that both introduce concealed hypotheses, and that the methods of the special theory cannot produce a unique answer.

Standard works on relativity still start from the postulated invariance of the velocity of light, which can be stated in the form

$$ds=0 \text{ is equivalent to } ds'=0 \dots\dots\dots (1)$$

for the observers, and infer that there is a linear relation between the coordinate systems, leading to the Lorentz-Einstein transformation. It was, I think, first shown by E. Cunningham that the conclusion does not follow; there are infinitely many relations that satisfy the equivalence, which do not even need to be linear. I have given additional conditions that are sufficient to lead to the transformation (Jeffreys 1957). The first is that for two observers of the same body

$$\frac{d^2x}{dt^2} = \frac{d^2y}{dt^2} = \frac{d^2z}{dt^2} = 0 \text{ is equivalent to } \frac{d^2x'}{dt'^2} = \frac{d^2y'}{dt'^2} = \frac{d^2z'}{dt'^2} = 0. \dots (2)$$

This amounts to saying that two observers will agree on what particles move with uniform velocity in straight lines. It does not say that there are no other

* Manuscript received June 30, 1958.

† St. John's College, Cambridge, England.

particles. Nor does it say that it holds for all pairs of observers; it, like (1), picks out a class of observers whose measures satisfy the rules. I shall call them unaccelerated observers. (1) and (2) together can be shown to imply

$$ds' = kds, \dots\dots\dots (3)$$

where k is a constant for a given pair of unaccelerated observers.

The second hypothesis is that

$$k=1. \dots\dots\dots (4)$$

This requires some actual comparison of scales. One postulate, experimentally verifiable in principle, that suffices to justify it is that observers will attach the same measures to displacements normal to their relative velocity. The essential point is that the Lorentz-Einstein relation between systems of reference needs hypotheses equivalent to Newton's first law and the establishment of some comparison of scales; it has no relevance to accelerated particles without further hypothesis.

In the conditions of the problem R is an unaccelerated particle. M is initially in contact with R and they have been together long enough to synchronize their clocks, so that $k=1$. R continues in its original path. M moves away from it in a straight line with uniform velocity. At some stage M rebounds from an obstacle, and ultimately meets R again, and is brought to relative rest; then the clocks are again compared. Builder argues that there is an asymmetry between R and M , Dingle that there is not. Now the conditions considered in the special theory contemplate a whole background of unaccelerated observers, who would agree that R is unaccelerated, and that M has three impulsive changes of velocity. (There is no objection to supposing R so massive in comparison with M that its two impulsive changes of velocity can be neglected.) It is not true, as Dingle supposes, that R would be the only authority for the change of direction of motion of M ; consequently his *a priori* argument for symmetry falls to the ground. But, since M is an accelerated particle, we cannot apply the special theory to it without further hypothesis, even though the accelerations are confined to the duration of the impulses.

Even if R and M were the only observers Dingle's argument still fails. He says "The principle of relativity allows us with equal justification to suppose that R is moving and M stationary. . . The two clocks will agree on reunion. That this must be so follows immediately from the symmetry of the situation and the principle of relativity of motion." It would be possible to infer a change of relative motion from a change in the Doppler effect; the two observers would then infer that *one* of them was accelerated, but could not say which. The symmetry is not one of fact but of knowledge. To make it into one of fact they would have to agree to attribute precisely half the change to each observer, but this is not what Dingle says. If a difference is to be expected at all, the observers would both expect that the clocks would differ in return, and that the sign of the difference would reveal which was on the accelerated body. Not to know which way the difference will be is not the same thing as to know that it is 0; this would be like saying that if $x^2=1$, then $x=\pm 1$, therefore $x=0$. This statement

is still too favourable to Dingle's point of view, since it supposes that the observers themselves have no means of detecting which has changed its velocity, before the return. Actually the reversal of M 's velocity implies an impulse; M would presumably have seen the body it collided with or at least felt the bump.

Suppose that in R 's system M proceeds out with velocity u_1 to a distance X and returns with velocity $-u_2$. There is no need for u_1 and u_2 to be equal, since the obstacle may be moving or restitution may be imperfect. Then in R 's system the total time is

$$T_R = X \left(\frac{1}{u_1} + \frac{1}{u_2} \right). \quad \dots\dots\dots (5)$$

In each part of the path M is in uniform motion. If $(xyzt)$ refer to R 's system, $(x'y'z't')$ to M 's, we have for the outward journey

$$ds' = k_1 ds, \quad dx/dt = u_1, \quad dx'/dt' = 0, \quad \dots\dots\dots (6)$$

whence

$$ds' = c dt' = k_1 (c^2 - u_1^2)^{\frac{1}{2}} dt. \quad \dots\dots\dots (7)$$

Similar relations hold for the return journey; then in M 's system the total time is

$$T_M = X \left\{ \frac{k_1}{u_1} \left(1 - \frac{u_1^2}{c^2} \right)^{\frac{1}{2}} + \frac{k_2}{u_2} \left(1 - \frac{u_2^2}{c^2} \right)^{\frac{1}{2}} \right\}. \quad \dots\dots\dots (8)$$

If $k = (1 - u^2/c^2)^{-\frac{1}{2}}$, $T_M = T_R$ and Dingle's result follows. If $k=1$, Builder's result follows. The question is which, if either, is right. The time τ taken by the impulses does not matter since it is small and anyhow presumably does not increase indefinitely with X .

$k=1$ in the special theory supposes some method of comparison of scales. Here M 's clock is compared with R 's before departure and $k=1$. But as M 's velocity changes the only obvious method of comparison after departure is by measures of transverse displacements by both. M could always take scales to keep $k=1$, but the question is whether, if he did, he would still assign the same length to the same scale. If there is a change, R could detect it. Suppose the linear dimensions of M are of order a , and its density ρ . Then, if it acquires velocity u in travelling through its linear dimensions, momentum per unit area ρau is acquired in time a/u , and the force needed per unit area is of order ρu^2 . Before the expulsion this is counterbalanced by other reactions, during the expulsion it is not. Then the stress implies elastic strains of order $\rho u^2/\mu$, where μ is some elastic constant of the material. If α is the velocity of elastic waves the strains are of order u^2/α^2 . Now the general theory makes the greatest possible α of order c ; actual values are much less than c . Hence the expulsion gives changes of length differing from 1 by far more than $(1 - u^2/c^2)^{\frac{1}{2}}$ does. This concerns lengths in the direction of motion, but transverse lengths will be altered in a comparable ratio by the Poisson's ratio effect.

The conclusion seems to be that Builder's result cannot be right in any case. Dingle's might be right if M is made of the most rigid material that is possible,

subject to the condition that the velocities of elastic waves do not exceed that of light, though his way of getting it is fallacious. With actual materials it also will be wrong. The change of k is of course one of the phenomena that could be covered by the general theory applied to elasticity.

References

- BUILDER, G. (1957).—*Aust. J. Phys.* **10**: 246–62.
 DINGLE, H. (1957).—*Aust. J. Phys.* **10**: 418–23.
 JEFFREYS, H. (1957).—“Scientific Inference.” Ch. 8. (Cambridge Univ. Press.)

THE CLOCK PARADOX IN RELATIVITY*

By E. F. FAHY†

I have been following with great interest the discussions on the above topic which have been published in several journals. On reading a recent letter to *Nature* by Professor Herbert Dingle (1957) in reply to a previous letter by Sir Charles Darwin (1957), I noticed that it is feasible to perform astronomical observations which could provide an experimental basis for choosing between the two points of view. In fact, these observations may have already been made.

Dingle (1957), in the course of his analysis of the particular aspect of this problem which was introduced by Darwin (1957), concludes that “ S_1 will not observe S_0 ’s flashes to change until after he has fired his rocket”. It is evident from the context that the length of the delay is

$$t_1 - \frac{1}{2}T_1 = \frac{1}{2}T_1(1 - \beta)/(1 + \beta).$$

In the light of this result, consider an observer on this Earth who is interested in one of the distant nebulae. He sees a red-shift in the spectral lines and can think of himself as being a traveller who left that nebula many years ago, thereby interpreting the red-shift in terms of the velocity which he believes he gave himself at the beginning of his journey. He now decides that he will return to the nebula and builds a rocket which will take him from the Earth and produce a violet-shift in the nebula’s spectrum equal in magnitude to the previous red-shift. Dingle’s result indicates that this traveller will have to wait for a long time before he will observe the violet-shift; if D light years is the distance to the nebula, it indicates that the delay is about $D/2$ years.

On advancing the argument further, one arrives at the following aspect of the above situation, which is simpler from the experimental point of view. Consider any star which lies approximately in the plane of the Earth’s orbit. Because of the Earth’s orbital motion, an earth-bound observer would expect to

* Manuscript received February 17, 1958.

† Physics Department, University College, Cork, Ireland.

find that any visible line in the spectrum of this star should oscillate about a mean position with a period of 1 year and an amplitude of approximately 0.5 \AA . Dingle's result in this case indicates, however, that the oscillation in the spectrum would be many revolutions out of phase with the Earth's orbital motion; in the case of most stars, the phase difference would not be an integral number of revolutions. An effect of this kind should be observable, since it means that the observed oscillations in the spectra of the stars are, in general, out of phase with one another.

A similar argument, based on Professor Dingle's point of view, would lead to the result that the phases of the apparent oscillatory motions of the stars due to aberration should, in general, vary from star to star.

In both cases (Döppler effect as well as aberration) the other point of view predicts that there should be no phase differences.

I think that it would be pertinent to this discussion if an astronomer would tell us whether or not such phase differences exist.

References

- DARWIN, C. (1957).—*Nature* **180**: 976.
DINGLE, H. (1957).—*Nature* **180**: 1275.

VARIATIONS IN IONOSPHERIC *F*-REGION CHARACTERISTICS*

By N. M. BRICE†

It is well known that *F*-region characteristics in the undisturbed ionosphere change with season and with sunspot activity and that these effects are greatest at high geomagnetic latitudes (Maeda 1955; Martyn 1955).

h'f records have been obtained at Macquarie Island (geomagnetic latitude 60°S ., local magnetic dip angle 78°) since 1950 and some of these were examined to study the above changes.

Method of Analysis

At the suggestion of Dr. D. F. Martyn, some Macquarie Island *h'f* records were scaled and electron density versus true height (*N, h*) curves computed on Silliac (School of Physics, University of Sydney) using the programme of Duncan (1958). Records were chosen near midday and midnight, summer and winter, and at sunspot maximum (S. max.) and sunspot minimum (S. min.) activity.

During sunspot minimum, no records were obtainable at night in winter, owing to very low critical frequencies, spread echoes, or noise.

* Manuscript received August 13, 1958.

† Physics Department, University of Queensland, Brisbane.

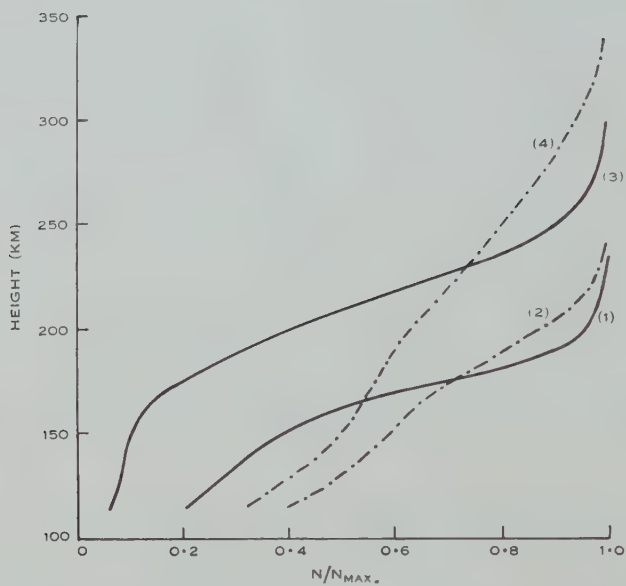


Fig. 1 (a).—Day. (1) Winter sunspot minimum, (2) summer sunspot minimum, (3) winter sunspot maximum, (4) summer sunspot maximum.

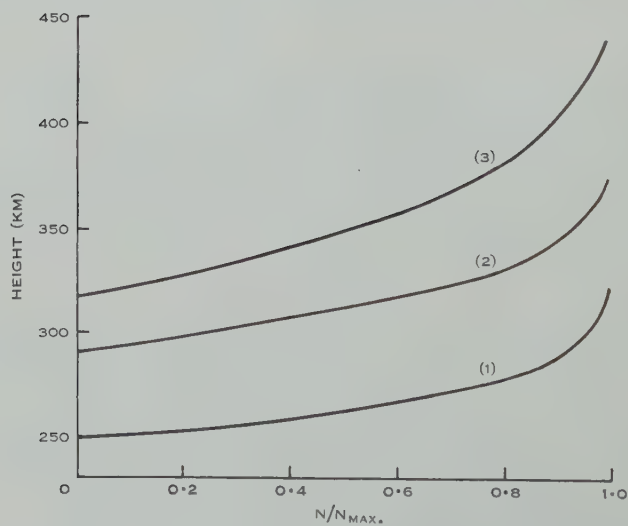


Fig. 1 (b).—Night. (1) Summer sunspot minimum, (2) winter sunspot maximum, (3) summer sunspot maximum.

Some uncertainty of true height occurred in the E region, as no frequencies less than 1.22 times the gyro frequency in the F region (i.e. below 2.05 Mc/s) could be used on this programme (Duncan 1958). However, the effect of this at F -region heights is negligible. It was assumed that no substantial trough in electron density occurred between layers, and records were restricted as far as possible to magnetically quiet days.

Between six and ten records were used for each period and a mean curve was obtained by plotting N/N_{\max} against true height for each record and averaging heights. According to Schmerling and Thomas (1956), this is sufficient to show the character of the (N, h) profile, and inspection of the curves showed a strong similarity of shape between the mean curve and the individual curves used in each period.

Results

Figure 1 shows the mean curves for the periods examined and Table 1 gives some parameters derived from the (N, h) curves.

In this presentation h_{\max} and N_{\max} refer to the height and electron density at the level of maximum ionization, and S. max. and S. min. to observations at maximum and minimum sunspot activity respectively.

Chief features of interest are:

(a) *Heights*.—(i) Heights of the maximum electron density of both F_1 and F_2 layers are greater at S. max. than S. min. (ii) h_{\max} is greater at night than during the day at all epochs (Mitra 1952). (iii) h_{\max} is greater in summer than in winter. By day this effect is more pronounced at S. max. than S. min.

(b) *Total Electron Content below N_{\max}* .—(i) The excess of electron content in summer over that in winter is due equally to increases in low height ionization (E and F_1 layers) and in F -region thickness and to a less extent to greater N_{\max} . (ii) At night (S. max.) the summer increase is due to greater N_{\max} and increased thickness of the layer. (iii) By day (S. max.) despite the summer increase in F_2 thickness and in E and F_1 ionization, the total electron content is greater in winter due to the greatly increased value of N_{\max} .

(c) *Scale Heights*.—The difference between h_{\max} and the height at $0.698N_{\max}$ (corresponding to one scale height of a Chapman layer) and between h_{\max} and the height at $0.958N_{\max}$ (or 0.4 scale heights) were found. From these it is seen that (i) by night the profiles are not incompatible with Chapman layers; (ii) by day in winter, the curvature of the profile approaching N_{\max} increases more rapidly than expected for a Chapman layer, whereas in summer it increases less rapidly than expected (at S. max.) or fits a Chapman layer (at S. min.).

Discussion

These curves are, to the best of the author's knowledge, the first high southern latitude (N, h) profiles calculated. While no attempt is made to discuss the implications of the results, it is felt that they will be of assistance to those working

TABLE 1
PARAMETERS DERIVED FROM THE (N, h) CURVES

Period	Average Height of F_1 Maximum Ionization (km)	Height F_2 (km)	F Region Thickness (km)	$N_{\max.}$ ($\times 10^6$ cm $^{-2}$)	Total Electron Content below $N_{\max.}$ ($\times 10^{12}$ cm $^{-2}$)	Chapman Scale Height at $Z=-1$ (km)	Chapman Scale Height at Nose (km)
Summer Sunspot Minimum Day ..	169	240	124	0.338	3.3	66	65
Winter Sunspot Minimum Day ..		235	115	0.328	2.7	60	90
Summer Sunspot Maximum Day ..	202	338	219	0.606	9.9	119	85
Winter Sunspot Maximum Day ..		298	175	1.772	16.4	73	85
Summer Sunspot Minimum Night ..		320	71	0.112	0.6	49	55
Summer Sunspot Maximum Night ..		437	121	0.456	3.3	72	66
Winter Sunspot Maximum Night ..		372	81	0.211	1.2	51	55

on the theory of the quiet ionosphere. They certainly substantiate King's statement (1957) that $h_p F_2^*$ is meaningless during the day.

Acknowledgment is due to the Ionospheric Prediction Service, Department of the Interior, for the loan of the records which were obtained in cooperation

* $h_p F_2^*$, defined as the virtual height at 0.834 of the critical frequency, is commonly used as the real height of $N_{\max.}$. This is only correct for a parabolic (N, h) profile.

with Antaretic Division, Department of External Affairs. The author wishes to express his thanks also to Dr. D. F. Martyn, F.R.S., and Mr. R. A. Duncan (Radio Research Laboratories) and Dr. J. A. Thomas (University of Queensland) for advice and assistance with this work and to Professor H. Messel (University of Sydney) for permission to use Silliac.

References

- DUNCAN, R. A. (1958).—*J. Geophys. Res.* **63**, No. 3.
 KING, G. A. M. (1957).—*J. Atmos. Terr. Phys.* **11**: 237.
 MAEDA, K. (1955).—"The Physics of the Ionosphere," p. 245. (Phys. Soc.: London.)
 MARTYN, D. F. (1955).—"The Physics of the Ionosphere," p. 254. (Phys. Soc.: London.)
 MITRA, S. K. (1952).—"The Upper Atmosphere," p. 297. (Asiatic Society: Calcutta.)
 SCHMERLING, E. R., and THOMAS, J. O. (1956).—*Phil. Trans. A* **248**: 609.

APPROXIMATIONS FOR THE ELECTRON DENSITY IN METEOR TRAILS*

By A. A. WEISS†

Herlofsen (1947) has shown that an exact solution of the equations governing the evaporation of a meteor particle during its flight in the upper atmosphere can be obtained in the special case of a spherical meteor in an isothermal atmosphere. In terms of the number n of meteor atoms evaporated in unit length of trail, the electron density α is

$$\alpha = \beta n. \quad \dots\dots\dots (1)$$

β is the ionizing probability, defined as the probability that an evaporated meteor atom will produce a free electron. If a power-law dependence of β upon velocity v is assumed, so that

$$\beta = \beta_0 v^\eta, \quad \dots\dots\dots (2)$$

the electron density may be written

$$\alpha = K(\eta) \left(\frac{v}{v_\infty} \right)^\eta \left(\frac{\rho H \sec \chi}{r_\infty \rho_m} \right) \left(\frac{v^2}{12l} \right) \left(\frac{r}{r_\infty} \right)^2, \quad \dots\dots\dots (3)$$

$$K(\eta) = 4 \cdot 5 \beta_0 m_\infty v_\infty^\eta \cos \chi / \mu H.$$

In these expressions m , r , and ρ_m are the mass, radius, and density of the meteor; μ the mass of an individual meteor atom; l the latent heat of evaporation of a meteor atom, corrected for the efficiency of heat transfer; ρ the density of the atmosphere and H the atmospheric scale height; and χ is the zenith angle of the meteor radiant. The subscript ∞ refers to the initial state of the meteor.

* Manuscript received June 23, 1958.

† Division of Radiophysics, C.S.I.R.O., at Department of Physics, University of Adelaide.

If it is assumed that the deceleration of the meteor is small over the whole flight, and that v is sufficiently large so that $v^2 \gg 12l$ and also $v^2 \gg 3l(2+\eta)$, the following expressions give a good approximation to the electron density profile:

$$\alpha = (\beta/\mu H) m_\infty \cos \chi (\rho/\rho_{\max.}) [1 - \frac{1}{3}(\rho/\rho_{\max.})]^2, \quad \dots (4a)$$

$$\rho_{\max.} = (8l/3v_\infty^2)(\rho_m r_\infty/H) \cos \chi. \quad \dots (4b)$$

These approximations have passed into general use, and a good deal of the theory underlying the interpretation of radio-echo observations of meteors has been based on them.

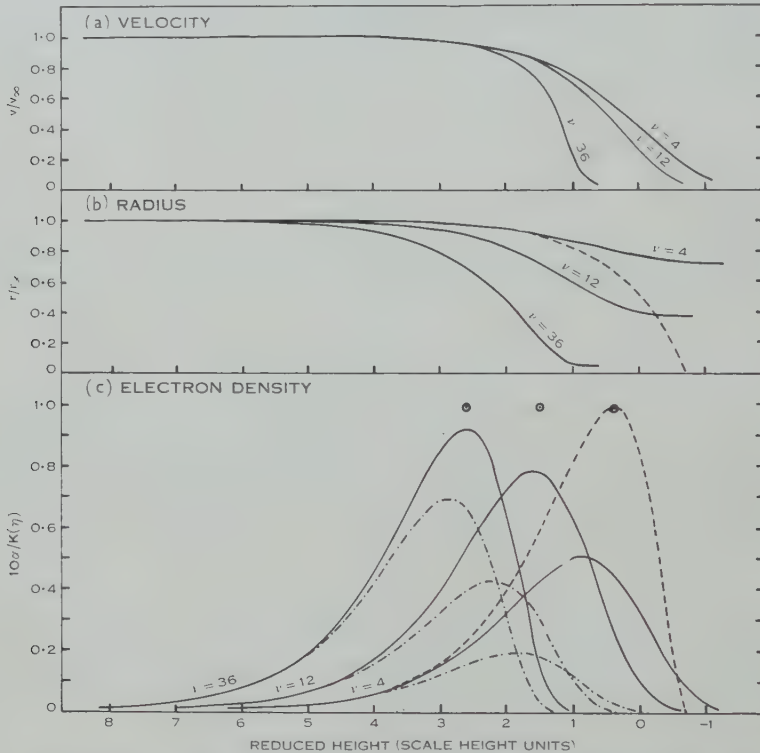


Fig. 1.—Comparison of exact and approximate solutions of the evaporation and ionization equations for a spherical meteor particle in an isothermal atmosphere. η =ionizing probability exponent of an evaporated meteor atom; $v=v_\infty^2/l$. See text for constant $K(\eta)$. Exact solution: $\eta=0$, ———; $\eta=5$, - - - - -. Usual approximation, all η , - · - · - ·. Maximum electron density for all η with old approximation, \odot . Velocity and radius are not functions of η .

Calculations of deceleration with Herlofson's exact solution and insertion of numerical values into the other two assumptions involving v , l , and η , show that all three assumptions are justified for fast meteors. The validity of their application to slow meteors is open to serious doubt unless $l < 10^{11}$ ergs/g and $\eta \sim 0$. Jacchia (1949) has given a mean value of $l \sim 3 \times 10^{11}$ ergs/g, and several independent evaluations of η (Whipple 1955; Hawkins 1956; Weiss 1957) suggest a value in the vicinity of 5. For slow meteors the divergence of the

approximate expressions (4) from the exact solution (3) therefore deserves careful consideration.

The electron density profiles and the velocity and radius of the meteor during its flight, as given by the exact solution (3) and the approximations (4), are compared in Figure 1 for three values of the parameter $v=v_{\infty}^2/l$, namely, $v=4, 12, 36$. To avoid confusion, only one set of approximate curves, for $v=4$, appears in this diagram. The shape of the approximate curves is independent of v and η . For other values of v the electron density curves may be located by reference to the circled dots which show the positions of the maxima. In Figure 1, and also Table 1, the reduced height is equivalent to $\ln (\rho H \sec \chi / \rho_m r_{\infty})$.

TABLE 1

COMPARISON OF APPROXIMATIONS WITH EXACT SOLUTION FOR POINT OF MAXIMUM ELECTRON DENSITY

v_{∞}^2/l	4	4	12	12	36	36
η	0	5	0	5	0	5
<hr/>									
(a) $10\alpha_{\max.}/K(\eta)^*$									
Exact solution	..			0.504	0.192	0.782	0.423	0.917	0.692
New approximation	..			0.772	0.220	0.939	0.467	0.980	0.726
Old approximation	..			0.988	0.988	0.988	0.988	0.988	0.988
<hr/>									
(b) Reduced Height*									
Exact solution	..			0.90	1.89	1.57	2.21	2.63	2.88
New approximation	..			1.10	1.91	1.79	2.27	2.71	2.93
Old approximation	..			0.39	0.39	1.50	1.50	2.60	2.60

* See text for $K(\eta)$ and for definition of reduced height.

The approximations (4) evidently give a poor representation of the exact solution for large values of η and small values of v_{∞}^2/l . A more realistic description of conditions near the point of maximum electron density is afforded by the following improved set of approximations:

$$\alpha_{\max.}=12(\beta_0 v_{\max.}^{\eta}/\mu H)m_{\infty}F(\eta)^2[1+2F(\eta)]^{-3}\cos\chi, \quad \dots (5a)$$

$$\rho_{\max.}=(8l/v_{\infty}^2)(\rho_m r_{\infty}/H)[1+2F(\eta)]^{-1}\cos\chi, \quad \dots (5b)$$

$$v_{\max.}^2=v_{\infty}^2+12l\ln(r_{\max.}/r_{\infty}), \quad \dots (5c)$$

$$r_{\max.}/r_{\infty}=2F(\eta)[1+2F(\eta)]^{-1}, \quad \dots (5d)$$

where

$$F(\eta)=1+3(2+\eta)l/v_{\infty}^2. \quad \dots (6)$$

In arriving at these approximations, the assumption $v^2\gg 3l(2+\eta)$ has been abandoned. Deceleration has been neglected in equations (5b), (5c), and (5d) relating to the evaporation process, but is introduced into (5a) for the electron density. The extent of the improvement over the old approximations may readily be seen from Table 1.

The shape of the electron density profile depends on the deceleration along the whole meteor path. Empirical improvements to the profile (4a), which

fails only over the last third of the meteor trail, could be suggested. Departures of the actual atmosphere from the assumed isothermal model and the effects of possible fragmentation of the meteor body (Jacchia 1955) will limit the value of further refinements in this direction. For the same reason, there is little point in pushing the approximations for the maximum point beyond the degree of accuracy attained in the set of approximations (5).

References

- HAWKINS, G. S. (1956).—*Astrophys. J.* **124**: 311.
HERLOFSEN, N. (1947).—*Rep. Progr. Phys.* **11**: 444.
JACCHIA, L. G. (1949).—H.C.O. and Cent. Anal. M.I.T. Tech. Rep. No. 3 (H.R. Ser. II—31).
JACCHIA, L. G. (1955).—*Astrophys. J.* **121**: 521.
WEISS, A. A. (1957).—*Aust. J. Phys.* **10**: 397.
WHIPPLE, F. L. (1955).—*Astrophys. J.* **121**: 241.

INDEX

	PAGE		PAGE
Aerials, Gain Measurements of, used in Interferometer and Cross-type Radio Telescopes	70	Builder, G.— Ether and Relativity ..	279
Allum, F. R.— <i>See</i> Spicer, B. M.	273, 298	The Constancy of the Velocity of Light	457
Argon, Liquid, Ionization by α -Particles	158	Bullen, K. E. — <i>See</i> Burke- Gaffney, T. N.	318
Atmosphere, Thermal Structures in the Lowest Layers of the ..	168	Burgess, V. R.— <i>See</i> Loughhead, R. E.	35
Aurora Australis, Low Latitude Reflections from the ..	126	Burke-Gaffney, T. N., and Bullen, K. E.— On the Seismological Aspects of the 1954 Hydrogen Bomb Explosions	318
Baglin, J. E.— <i>See</i> Spicer, B. M.	273, 298	Centaurus, Strong Radio Sources in	400
Bailey, G. M., and Prescott, J. R.— Decay Time of the Lumines- cence of a Zinc Sulphide Neutron Detector for Neutron and γ -Ray Excita- tion	135	Centaurus-A, The Radio Emis- sion from	517
Beryllium, Excited States of ^8Be from the $^7\text{Li}(d,n)^8\text{Be}$ Re- action	502	Clock Paradox in Relativity ..	586
Blevin, H. A., and Haydon, S. C.— The Townsend Ionization Co- efficients in Crossed Electric and Magnetic Fields ..	18	Clock Paradox in Special Re- lativity	583
Bolie, V. W.— Electromagnetic Propagation in an Almost Homogeneous Medium	118	Conductivity, Thermal, on the Cylindrical Probe Method of Measuring the ..	255, 409
Bowen, E. G.— Freezing Nucleus Measure- ments in January 1957 ..	452	Corona, Solar, Evidence of Echoes in the, from a New Type of Radio Burst ..	215
Bray, R. J.— <i>See</i> Loughhead, R. E.	177	Coulomb Wave Functions ..	138
Bray, R. J., and Loughhead, R. E.— Observations of Changes in the Photospheric Granules ..	507	Counter, Type I, Methods for Numerical Calculations with the	143
The Lifetime of Sunspot Fila- ments	185	Coupling of Nuclear Spins in Molecules	1
Brice, N. M.— Variations in Ionospheric F - Region Characteristics ..	587	Crisp, R. S.— The Soft X-ray L_{23} Emission Spectrum of Magnesium from Solid and Evaporated Targets	449
		Cruikshank, F. D.— The Design of Photographic Objectives of the Triplet Family. I. The Design of the Triplet Type III Ob- jective	41

	PAGE		PAGE
Decay Time of the Luminescence of a Zinc Sulphide Neutron Detector for Neutron and γ -Ray Excitation	135	Galvanomagnetic Effect in Isotropic Metals and Semi- conductors	235
Elastic Constants of Isotropic Materials in n Dimensions, The General Relationships between the	154	Gardner, F. F., and Shain, C. A.— Further Observations of Radio Emission from the Planet Jupiter	55
Electromagnetic Propagation in an Almost Homogeneous Medium	118	Gazzard, A. D.— Lunar Tides in E_2 , at Brisbane	272
Electron Density in Meteor Trails, Approximations for..	591	Giovanelli, R. G.— Flare-Puffs as a Cause of Type III Radio Bursts ..	350
Electron Excitation of Collective Nuclear Transitions	481	Giovanelli, R. G., and McCabe, Marie K.— The Distribution of Flare Heights as derived from Limb Flares	130
Electrons, Electromagnetic Radiation from, Rotating in an Ionized Medium under the Action of a Uniform Magnetic Field	424	The Flare-surge Event ..	191
Electrons, Ferromagnetic Ex- change between Coupled Pairs of	447	Giovanelli, R. G., and Roberts, J. A.— Optical Observations of the Solar Disturbances causing Type II Radio Bursts ..	353
Elford, W. G.— <i>See</i> Mainstone, J. S.	277	Granules, Photospheric, Obser- vations of Changes in the ..	507
Ellyett, C. D.— <i>See</i> Seed, T. J...	126	Guest, P. G.— Methods for Numerical Calcu- lations with the Type I Counter	143
Ether and Relativity	279	Haydon, S. C.— <i>See</i> Blevin, H. A.	18
Fahy, E. F.— The Clock Paradox in Re- lativity	586	Heisler, L. H.— Anomalies in Ionosonde Re- cords due to Travelling Ionospheric Disturbances	79
Ferromagnet, Fluctuating Field, at Low Temperatures ..	310	Hill, E. R.— A Search for Radio Emission at 3.5 m from the Local Supergalaxy	580
Ferromagnetic Exchange be- tween Coupled Pairs of Electrons	447	<i>See also</i> Mills, B. Y. ..	360
Flare-Puffs as a Cause of Type III Radio Bursts	350	Hill, E. R., Slee, O. B., and Mills, B. Y.— A Pencil-beam Survey of the Galactic Plane at 3.5 m ..	530
Flares, Solar, Distribution of Flare Heights as Derived from Limb Flares	130	Hydrogen Bomb Explosions, Seismological Aspects of the 1954	318
Flare-surge Event, The.. ..	191	Ionization of Liquid Argon by α -Particles	158
Fornax, Strong Radio Sources in	400		
Fornax-A, The Radio Emission from	517		
Gain Measurements of Aerials used in Interferometer and Cross-type Radio Telescopes	70		

PAGE

Ionization, Townsend Coefficients in Crossed Electric and Magnetic Fields ..	18
Ionosphere, Lunar Tides in E_{2s} at Brisbane	272
Ionosphere, "Spread- F " Ionospheric Echoes at Night at Brisbane	322
Ionosphere, Variations in F -Region Characteristics ..	587
Ionospheric Disturbances, Travelling, Anomalies in Ionosonde Records due to ..	79
Ionospheric Disturbances, Travelling, in the F Region..	91
Jeffreys, H.— The Clock Paradox in Special Relativity	583
Jupiter, Further Observations of Radio Emission from the Planet	55
Komesaroff, M.— Polarization Measurements of the Three Spectral Types of Solar Radio Burst ..	201
Learner, A., and Robson, B. A.— Coulomb Wave Functions ..	138
Light, Velocity of, The Constancy of the.. ..	457
Little, A. G.— Gain Measurements of Large Aerials used in Interferometer and Cross-type Radio Telescopes	70
Loughhead, R. E.— <i>See</i> Bray, R. J.	185, 507
Loughhead, R. E., and Bray, R. J.— The Wilson Effect in Sunspots	177
Loughhead, R. E., and Burgess, V. R.— High Resolution Cinematography of the Solar Photosphere	35
Luminescence of a Zinc Sulphide Neutron Detector, Decay Time of, for Neutron and γ -Ray Excitation	135

PAGE

McCabe, Marie K. — <i>See</i> Giovanelli, R. G. ..	130, 191
Magnesium, Photoneutrons from Natural	273
Magnesium, Soft X-Ray L_{23} Emission Spectrum of ..	449
Mainstone, J. S., Elford, W. G., and Weiss, A. A.— A Suggested Improvement to the C.W. Technique for Measurement of Meteor Velocities	277
Metals, Isotropic, On the Galvanomagnetic, Thermomagnetic and Thermoelectric Effects in	235
Meteor Shower, The 1956 Phoenicid	113
Meteor Trails, Approximations for the Electron Density in..	591
Meteor Velocities, A Suggested Improvement to the C.W. Technique for Measurement of	277
Mills, B. Y.— <i>See</i> Hill, E. R. ..	530
Mills, B. Y., Slee, O. B., and Hill, E. R.— A Catalogue of Radio Sources between Declinations $+10^\circ$ and -20°	360
Moore, E. J.— On the Galvanomagnetic, Thermomagnetic, and Thermoelectric Effects in Isotropic Metals and Semiconductors	235
Munro, G. H.— Travelling Ionospheric Disturbances in the F Region	91
Nebulae, Hydrogen, On the Radio Emission of	388
Neutron Detector, Decay Time of the Luminescence of a Zinc Sulphide, for Neutron and γ -Ray Excitation	135
Neutrons, Photo-, from Natural Magnesium	273
Nuclear Transitions, Electron Excitation of Collective ..	481
Nuclei, Atmospheric Freezing, Measurements in January 1957	452

	PAGE		PAGE
Nuclei with Z between 9 and 30, Photodisintegration of ..	490	Radiation, Electromagnetic, from Electrons Rotating in an Ionized Medium under the Action of a Uniform Magnetic Field	424
Objectives, Photographic, of the Triplet Family, Design of. I..	41	Radiation Transfer and the Possibility of Negative Ab- sorption in Radio Astronomy	564
Parthasarathy, R.— <i>See</i> Swarup, G.	338	Radio Astronomy, Radiation Transfer and the Possibility of Negative Absorption in ..	564
α -Particles, Ionization of Liquid Argon by	158	Radio Burst, Evidence of Echoes in the Solar Corona from a New Type of	215
Peck, A. J.— <i>See</i> de Vries, D. A. 255, 409		Radio Burst, Solar, Polarization Measurements of the Three Spectral Types of	201
Photodisintegration of Nuclei with Z between 9 and 30 ..	490	Radio Bursts, Type II, Optical Observations of the Solar Disturbances causing ..	353
Photodisintegration of Tantalum, The Giant Resonance of the..	298	Radio Bursts, Type III, Flare- Puffs as a Cause of ..	350
Photoneutrons from Natural Magnesium	273	Radio Emission from the Planet Jupiter, Further Observations of	55
Photosphere, Solar High Resolu- tion Cinematography of the..	35	Radio Telescopes, Gain Measure- ments of Aerials used in Interferometer and Cross-type	70
Posener, D. W.— Coupling of Nuclear Spins in Molecules	1	γ -Ray Excitation, Decay Time of the Luminescence of a Zinc Sulphide Neutron De- tector for Neutron and ..	135
Prescott, J. R.— <i>See</i> Bailey, G. M.	135	Relativity, Ether and	279
Puppis, Strong Radio Sources in	400	Relativity, Special, The Clock Paradox in	583
Radiation, Cosmic Radio, A Pencil-beam Survey of the Galactic Plane at 3.5 m ..	530	Relativity, The Clock Paradox in	586
Radiation, Cosmic Radio, A Search for Radio Emission at 3.5 m from the Local Super- galaxy	580	Relativity, The Constancy of the Velocity of Light ..	457
Radiation, Cosmic Radio, Catalogue of Radio Sources between Declinations $+10^\circ$ and -20°	360	Rishbeth, H.— Radio Emission from the Vela- Puppis Region	550
Radiation, Cosmic Radio, Emission from Centaurus-A and Fornax-A	517	Roberts, J. A.— Evidence of Echoes in the Solar Corona from a New Type of Radio Burst ..	215
Radiation, Cosmic Radio, Emission from the Vela- Puppis Region	550	<i>See also</i> Giovanelli, R. G. ..	353
Radiation, Cosmic Radio, On The Radio Emission of Hydrogen Nebulae	388	<i>See also</i> Twiss, R. Q. ..	424
Radiation, Cosmic Radio, Strong Radio Sources in Centaurus, Fornax, and Puppis	400	Robson, B. A.— <i>See</i> Learner, A.	138

	PAGE		PAGE
Seed, T. J., and Ellyett, C. D.— Low Latitude Reflections from the Aurora Australis ..	126	Sun, Flare-Puffs as a Cause of Type III Radio Bursts ..	350
Seismological Aspects of the 1954 Hydrogen Bomb Ex- plosions	318	Sun, High Resolution Cinemato- graphy of the Solar Photo- sphere	35
Semiconductors, On the Gal- vanomagnetic, Thermomag- netic, and Thermoelectric Effects in	235	Sun, Observations of Changes in the Photospheric Granules ..	507
Shain, C. A.— The Radio Emission from Centaurus-A and Fornax-A	517	Sun, Optical Observations of the Solar Disturbances caus- ing Type II Radio Bursts ..	353
<i>See also</i> Gardner, F. F. ..	55	Sun, Solar Brightness Distribu- tion at a Wavelength of 60 Centimetres	338
Sheridan, K. V.— An Investigation of the Strong Radio Sources in Centaurus, Fornax, and Puppis ..	400	Sun, The Distribution of Flare Heights as derived from Limb Flares	130
Slee, O. B.— <i>See</i> Hill, E. R. ..	530	Sunspot Filaments, The Life- time of	185
<i>See also</i> Mills, B. Y. ..	360	Sunspots, The Wilson Effect in	177
Soils, Cylindrical Probe Method of Measuring the Thermal Conductivity of ..	255, 409	Surge, The Flare-surge Event..	191
Spear, R. H.— Excited States of ^8Be from the $^7\text{Li}(d,n)^8\text{Be}$ Reaction..	502	Swarup, G., and Parthasarathy, R.— Solar Brightness Distribution at a Wavelength of 60 Centi- metres. II. Localized Radio Bright Regions	338
Spicer, B. M.— The Photodisintegration of Nuclei with Z between 9 and 30	490	Tantalum, The Giant Resonance of the Photodisintegration of	298
Spicer, B. M., Allum, F. R., Baglin, J. E. E., and Thies, H. H.— Photoneutrons from Natural Magnesium	273	Tassie, L. J.— Electron Excitation of Col- lective Nuclear Transitions	481
Spicer, B. M., Thies, H. H., Baglin, J. E., and Allum, F. R.— The Giant Resonance of Photodisintegration of Tan- talum	298	Taylor, R. J.— Thermal Structures in the Lowest Layers of the Atmos- phere	168
Spins, Nuclear, Coupling of, in Molecules	1	Thermoelectric Effect in Iso- tropic Metals and Semicon- ductors	235
Stacey, F. D.— Ferromagnetic Exchange be- tween Coupled Pairs of Electrons	447	Thermomagnetic Effect in Iso- tropic Metals and Semicon- ductors	235
The Fluctuating Field Ferro- magnet at Low Temper- atures	310	Thies, H. H.— <i>See</i> Spicer, B. M. 273, 298	
The Ionization of Liquid Argon by α -Particles ..	158	Townsend Ionization Coefficients in Crossed Electric and Mag- netic Fields	18
		Tschoegl, N. W.— The General Relationships between the Elastic Con- stants of Isotropic Materials in n Dimensions	154

	PAGE		PAGE
Twiss, R. Q.—		Wade, C. M.—	
Radiation Transfer and the		On the Radio Emission of	
Possibility of Negative Ab-		Hydrogen Nebulae ..	388
sorption in Radio Astronomy	564	Wave Functions, Coulomb ..	138
Twiss, R. Q., and Roberts, J. A.—		Webster, H. C.—	
Electromagnetic Radiation		A Study of "Spread- <i>F</i> " Iono-	
from Electrons rotating in		spheric Echoes at Night at	
an Ionized Medium under		Brisbane. IV. Range	
the Action of a Uniform		Spreading	322
Magnetic Field	424	Weiss, A. A.—	
Vela-Puppis Region, Radio		Approximations for the	
Emission from the	550	Electron Density in Meteor	
de Vries, D. A., and Peck, A. J.—		Trails	591
On the Cylindrical Probe		The 1956 Phoenicid Meteor	
Method of Measuring		Shower	113
Thermal Conductivity with		<i>See also</i> Mainstone, J. S. ..	277
Special Reference to Soils—		Wilson Effect in Sunspots ..	177
I. Extension of Theory		X-ray L_{23} Emission Spectrum of	
and Discussion of		Magnesium	449
Probe Characteristics	255	Zinc Sulphide Neutron De-	
II. Analysis of Moisture		tector, Decay Time of the	
Effects	409	Luminescence of α , for Neutron	
		and γ -Ray Excitation ..	135



3 8198 303 614 521

THE UNIVERSITY OF ILLINOIS AT CHICAGO

